

# Estrategias de clasificación de texturas en imágenes forestales hemisféricas

Máster en Investigación en Informática

Curso 2009-2010

Proyecto de Fin de Máster



Departamento de Ingeniería del Software e Inteligencia Artificial

Facultad de Informática.

Universidad Complutense de Madrid.



Autor:  
Pablo Iglesias Marcos

Dirigido por:  
Gonzalo Pajares Martinsanz  
Pedro Javier Herrera Caro



# Autorización

El abajo firmante, matriculado en el Máster en Investigación en Informática de la Facultad de Informática, autoriza a la Universidad Complutense de Madrid (UCM) a difundir y utilizar con fines académicos, no comerciales y mencionando expresamente a su autor el presente Trabajo Fin de Máster: “Estrategias de clasificación de texturas en imágenes forestales hemisféricas”, realizado durante el curso académico 2009-2010 bajo la dirección de Gonzalo Pajares Martisanz y Pedro Javier Herrera Caro en el Departamento de Ingeniería del Software e Inteligencia Artificial y a la Biblioteca de la UCM a depositarlo en el Archivo Institucional E-Prints Complutense con el objeto de incrementar la difusión, uso e impacto del trabajo en Internet y garantizar su preservación y acceso a largo plazo.

Firmado:

Pablo Iglesias Marcos.

# Agradecimientos

En primer lugar quería agradecer a mis directores su apoyo en los momentos complicados, y por haberme guiado del modo en que lo han hecho para así llegar a buen puerto con el proyecto.

También quería dar las gracias a Ana, por saber entender mi modo de hacer las cosas y por su paciencia, ánimos y haber estado siempre ahí. Del mismo modo, darle las gracias por haberme hecho la vida más fácil y habernos dejado reunirnos en tu casa siempre que lo hemos necesitado.

Por último pero no menos importante, a mis padres. Por haberme dado la posibilidad de haber llegado hasta aquí, y por el interés que han demostrado a lo largo de la marcha del proyecto.

# Resumen del proyecto

## Resumen

El trabajo de investigación desarrollado tiene su origen en el interés suscitado por el Centro de Investigación Forestal (CIFOR) del Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) para automatizar los procesos de medición para inventarios forestales, mediante un dispositivo estereoscópico de captura de imágenes equipado con una lente de “ojo de pez”.

Entre otras tareas, los inventarios forestales tienen como finalidad la obtención del volumen de madera mediante medidas en los troncos de los árboles. Como paso previo al proceso de visión estereoscópica, una tarea importante consiste en identificar las texturas de los troncos presentes en las imágenes, en este caso hemisféricas.

El núcleo central del trabajo de investigación que se presenta ha consistido en abordar la problemática de la clasificación de texturas naturales en estas imágenes hemisféricas de exterior desde la perspectiva de la combinación de clasificadores clásicos, que por otra parte constituye una de las principales líneas de investigación abiertas actualmente en el área, tal y como se pone de manifiesto en la literatura especializada.

Se plantea la necesidad de automatizar el proceso de clasificación consiguiendo una estrategia no supervisada a partir de clasificadores de naturaleza supervisada combinados. El objetivo es mejorar los resultados de los clasificadores simples mediante la combinación.

## Palabras clave

Texturas en imágenes, clasificación de imágenes forestales, clasificador híbrido, agrupamiento borroso, clasificador de Bayes, mapas auto-organizativos, ventana de Parzen, algoritmo de Lloyd, teoría de la decisión multicriterio fuzzy.

## **Abstract**

This work stems from the interest generated by the Forest Research Centre (CIFOR) part of the National Institute for Agriculture and Food Research and Technology (INIA) to automate the process of extracting information through a measurement mechanism for forest inventories. This mechanism is a stereoscopic image-based system equipped with fish-eye lenses.

One inventory task is the computation of wood volume through measurements from the tree trunks. As a previous step to the stereovision process, an important task consists in the texture identification belonging to the trunks in the hemispheric images.

This research work focuses on the classification of textures in this kind of outdoor images oriented toward the trunk identification. This is carried out by combining classifiers, which is a current open research line in classification problems, as reported in the specialized literature.

The main goal is the automation of the classification process with the aim of achieving an unsupervised strategy from supervised classifiers under combination. The aim is to improve the classification results obtained by the simple classifiers through the combined approach.

## **Keywords**

Image textures, forest image classification, hybrid classifier, Fuzzy Clustering, Bayes Classifier, self-organizing maps, Parzen's windows, Lloyd algorithm, Multicriteria Decision Making Fuzzy.

# Índice

<b>Capítulo 1. Introducción</b>	1
1.1 Antecedentes	1
1.2 Identificación del problema	8
1.3 Motivación	11
1.4 Objetivos de la investigación	13
1.5 Aportaciones a la investigación	13
1.6 Organización de la memoria	14
 <b>Capítulo 2. Estado del arte</b>	15
2.1 Visión general	15
2.2 Métodos de clasificación de texturas en imágenes	19
2.3 Métodos de clasificación combinados	24
2.3.1 Clasificadores combinados de índole general	24
2.3.2 Combinación de clasificadores en texturas de imágenes	25
2.4 Características de las imágenes en la clasificación de texturas naturales	28
2.4.1 Características espectrales a nivel de píxel	29
2.4.2 Características con patrones espaciales	31
2.4.3 Combinación de características espectrales y espaciales	34
2.5 Métodos de clasificación mediante información espacial	35
2.6 Resumen del análisis bibliográfico	36

<b>Capítulo 3. Estrategia combinada de clasificación</b>	<b>39</b>
3.1 Introducción	39
3.1.1 Fases de entrenamiento y decisión	40
3.2 Descripción de los clasificadores individuales seleccionados	43
3.2.1 Agrupamiento borroso	43
3.2.1.1 Fase de entrenamiento	43
3.2.1.2 Fase de decisión	48
3.2.2 Clasificador Paramétrico: Bayesiano	48
3.2.2.1 Fase de entrenamiento	48
3.2.2.2 Fase de decisión	52
3.2.3 Estimación no paramétrica: ventana de Parzen	53
3.2.3.1 Fase de entrenamiento	53
3.2.3.2 Fase de decisión	57
3.2.4 Algoritmo Generalizado de Lloyd	58
3.2.4.1 Fase de entrenamiento	58
3.2.4.2 Fase de decisión	60
3.2.5 Mapas Auto-Organizativos	61
3.2.5.1 Fase de entrenamiento	61
3.2.5.2 Fase de decisión	65
3.3 Funciones criterio para la partición	65
3.3.1 Comportamiento de las funciones criterio	68
3.4 Método de clasificación combinado: MCDMF	69
3.4.1 Fase de entrenamiento	71



3.4.2	Descripción general del clasificador combinado	72
3.4.2.1	Formulación del problema y definiciones	72
3.4.2.2	Normalización	75
3.4.2.3	Proceso de decisión	75
3.4.3	Fase de decisión: combinación de clasificadores	77
<b>Capítulo 4.</b>	<b>Análisis de resultados</b>	<b>83</b>
4.1	Introducción	83
4.2	Métodos de evaluación de resultados	83
4.2.1	Validación cruzada (Cross-validation)	83
4.2.2	Imágenes de referencia o “Ground truth”	85
4.2.3	Criterio de calidad de los porcentajes	88
4.3	Características de las imágenes tratadas	89
4.3.1	Iluminación	89
4.3.2	Complejidad de las texturas	90
4.3.3	Preprocesado de las imágenes	92
4.3.4	Determinación del número inicial de clases	95
4.4	Resultados de los algoritmos individuales	96
4.4.1	Agrupamiento Borroso	97
4.4.2	Estimación no paramétrica: ventana de Parzen	98
4.4.3	Clasificador Paramétrico: Bayesiano	98
4.4.4	Algoritmo Generalizado de Lloyd	98
4.4.5	Mapas Auto-Organizativos	99
4.5	Resultados del clasificador híbrido	100

<b>Capítulo 5. Conclusiones y Trabajo futuro</b>	103
5.1 Introducción	103
5.2 Conclusiones	103
5.3 Trabajo futuro	104
 <b>Bibliografía</b>	 107

# Capítulo 1

## Introducción

### 1.1 Antecedentes

La tecnología de los dispositivos de captura de imágenes es cada vez más sofisticada, lo que provoca un mayor interés por la utilización de las imágenes en diferentes aplicaciones. De hecho, hasta no hace muchos años, eran relativamente escasas las aplicaciones basadas en imágenes, en concreto en el mundo empresarial. Actualmente, esta situación se está invirtiendo. Cada vez es mayor el número de aplicaciones basadas en Visión Artificial. Existen diversas áreas de utilización: Agricultura, Forestal, Conservación medioambiental, Meteorología, Detección de edificaciones o infraestructuras. Algunas de tales aplicaciones son las siguientes:

- Control de cultivos en agricultura, propiciado por la necesidad de conocer los cultivos programados para la recepción de subvenciones u otros aspectos relacionados.
- Cálculo y medición de parcelas agrícolas y tipo de cultivos a los que se dedica.
- Control de riegos agrícolas.
- Agricultura de precisión: aplicación de herbicida de forma selectiva en el tratamiento de malas hierbas evitando la contaminación medioambiental y reduciendo costes de producción.
- Evaluación de catástrofes naturales: fuegos, daños por inundaciones, heladas en cultivos agrícolas, nevadas, etc.
- Detección de cambios en determinadas zonas, principalmente urbanas para el control de edificaciones o impactos medioambientales.
- Vigilancia en prevención de catástrofes, por ejemplo fuegos o inundaciones.

- Control de fenómenos meteorológicos como es el retroceso de determinadas playas.
- Vigilancia: forestal, marítima.
- Detección de infraestructuras: carreteras, caminos forestales, cañadas reales, etc.
- Finalmente, destacar la aplicación objeto de esta memoria, que en sí misma constituye una aplicación novedosa al estar relacionada con los tratamientos de imágenes en el ámbito forestal y su aplicación de cara a la obtención de inventarios forestales, tales como obtención del volumen de madera en un área de bosque concreta o determinación del estado de crecimiento de los árboles, entre otras muchas actividades relacionadas en este sentido.

De lo expuesto previamente se puede deducir fácilmente la enorme cantidad de información que los sistemas de captura de imágenes pueden proporcionar. En concreto, la clasificación de las imágenes según sus texturas ha sido un campo muy estudiado desde hace mucho tiempo. Su objetivo consiste en realizar una partición de la imagen en varias regiones disjuntas, de tal forma que cada región se debe caracterizar por su homogeneidad con respecto a alguna medida previamente establecida. Este es el paso previo para abordar tareas dotadas con un mayor nivel de inteligencia y conocimiento como puede ser el reconocimiento de objetos, la comprensión de la imagen o la descripción de la escena (González y Woods, 2007; Pajares y Cruz, 2007).

Un ejemplo de clasificación clásico sería el que se muestra en la figura 1.1, consistente en la identificación de texturas naturales procedentes de sensores aerotransportados, cuya finalidad última está encaminada hacia aplicaciones concretas, entre las que destacan la ordenación del territorio o la prevención de posibles catástrofes, tales como incendios forestales. La figura 1.1(a) muestra una imagen original aérea del terreno. La figura 1.1(b) se corresponde con la clasificación de texturas realizada mediante el clasificador paramétrico Bayesiano, que se describe en el capítulo tres, por ser uno de los métodos utilizados en este trabajo. Esta figura ya da idea de los resultados derivados de la clasificación en este tipo de texturas.

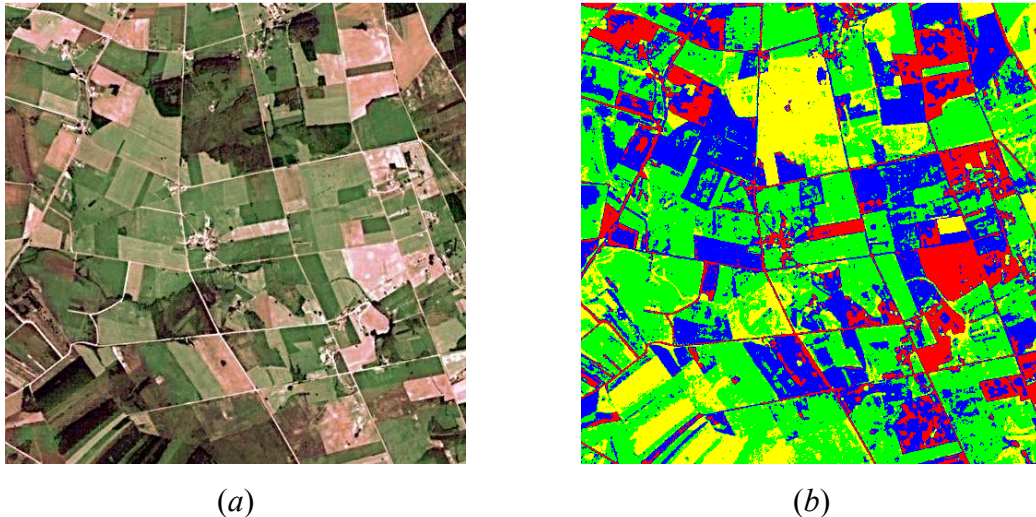


Figura 1.1 (a) Imagen aérea original; (b) identificación de las texturas naturales.

Resulta fácil deducir, a partir de la aplicación ilustrada en la figura anterior, la importante relevancia que establece el tratamiento de las texturas naturales. Se plantea entonces el problema de la clasificación de las mismas por medio del mejor procedimiento posible. De hecho, siguiendo ésta misma línea, diversos organismos y empresas desarrollan o utilizan aplicaciones para abordar esta problemática. Por citar algunos de ellos:

1. Dimap (2009), empresa dedicada en exclusiva al tratamiento de imágenes.
2. Proespacio (2009) agrupación de empresas del sector aeroespacial donde una de las actividades destacables es el desarrollo de aplicaciones mediante el uso de imágenes de satélite. En este consorcio destacan algunas empresas líderes del sector tanto en España como a nivel internacional: EADS-Espacio, EADS Astrium, CRISA, GMV, Indra Espacio, Sener, Hispasat, IberEspacio, Inasmet, Insa, Mier, NTE, Tecnológica, Ryma, Hispasat, GTD, Alcatel, CRISA, GTD.
3. Organismos oficiales y centros de investigación con los que existen trabajos de colaboración previos por parte de miembros del grupo ISCAR (2006), entre los que se encuentran los directores de este trabajo de investigación:
  - a. Instituto de Automática Industrial (IAI) del Consejo Superior de Investigaciones Científicas (CSIC) (Pajares y col., 2007; Tellaeche y col., 2008).
  - b. Instituto Nacional de Técnica Aeroespacial (INTA) (Pajares, 2006).

- c. Centro de Estudios y Experimentación de Obras Públicas (CEDEX), (Pajares y col. 2001, 2002).

La mayoría de las empresas citadas y algún organismo también mencionado, utilizan para desarrollar sus aplicaciones herramientas comerciales tales como las relacionadas a continuación:

- ERDAS Imagine (2009)
- Paquete de productos de información geográfica de Intergraph (2009)
- ENVI (2009) en lenguaje IDL (Canty, 2006)
- PCI Geomatics (2009)
- E-Cognition y Definiens (2009)
- Image Processing Toolbox de MATLAB (2009)

En general dichas herramientas poseen módulos de clasificación supervisada, de forma que requieren la intervención del usuario, quien ha de seleccionar y suministrar las muestras necesarias para el entrenamiento. Estos módulos suelen estar basados en clasificadores clásicos entre los que se encuentran los relacionados a continuación (Pajares y Cruz, 2007):

- Clasificadores estadísticos Bayesianos: máxima verosimilitud, ventana de Parzen.
- Agrupamiento borroso ("*Fuzzy Clustering*").
- Algoritmos competitivos: Cuantización vectorial ("*Learning Vector Quantization*"), Mapas Auto-Organizativos ("*Self-Organizing Feature Maps*").
- Red neuronal retropropagación ("*backpropagation*").

Excepto la red neuronal y la cuantización vectorial, los clasificadores anteriores se describen en el capítulo tres por ser la base para la combinación de los métodos propuestos en este trabajo de investigación.

Aunque la potencialidad de las herramientas citadas es evidente, los retos tecnológicos derivados de las aplicaciones mencionadas anteriormente hacen que en algunos casos el uso de tales herramientas resulte insuficiente para abordar con éxito las propuestas demandadas por los clientes según sus necesidades reales. E incluso, siendo suficientes, siempre cabe la posibilidad de introducir mejoras o modificaciones precisas según las necesidades, como ocurre realmente en numerosas aplicaciones. De esta manera, la *clasificación de texturas naturales* se perfila como uno de los procedimientos susceptibles de mejora, porque en la mayoría de los casos, los métodos simples de clasificación no proporcionan los resultados deseados, en concreto en imágenes de exterior como las utilizadas en este trabajo, cuya base son las texturas naturales. Surge por tanto, la necesidad de investigar nuevas alternativas en el ámbito de la clasificación, orientadas hacia la identificación de tales texturas.

Las imágenes empleadas en este trabajo son obtenidas a través del dispositivo de medición de bosques MU-200501738 patentado en 2005 por el Centro de Investigación Forestal (CIFOR) integrado en el Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA). La empresa “Consultoría Ferroviaria y Medioambiental, S.L.” desarrolló para el INIA un prototipo del dispositivo de medición, adaptado para una cámara digital Nikon® Coolpix® 4500 con una lente de “*ojo de pez*” del tipo FC-E8. En la figura 1.2(a) y (b) se muestra un par de imágenes estereoscópicas obtenidas mediante el dispositivo de medición mencionado previamente.

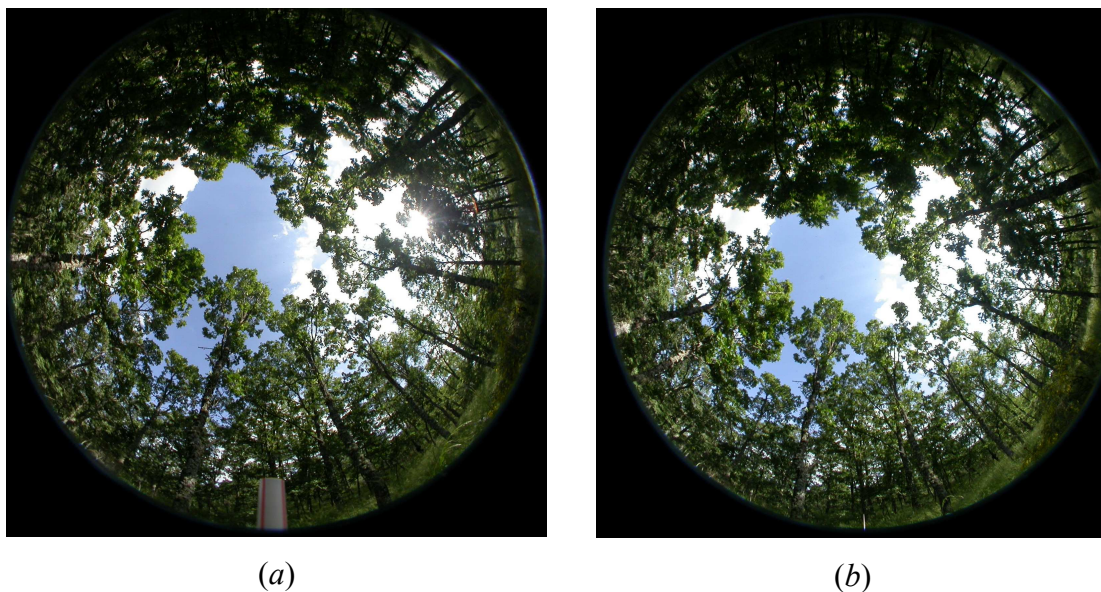


Figura 1.2 (a) y (b) Par estereoscópico obtenido mediante el dispositivo de medición MU-200501738.

Este dispositivo tiene como fundamento la *visión estereoscópica*, ya que permite obtener dos imágenes ligeramente desplazadas de la misma escena, esto es, una serie de parcelas representativas de una mayor extensión forestal a partir de las cuales se desea realizar *inventarios forestales*. En el inventario forestal se hacen mediciones con el fin de determinar diversos parámetros sobre la estructura poblacional y el estado de la masa forestal y cuantificar las existencias de madera, fruto, resina u otros productos forestales. Las parcelas suelen replantearse distribuidas de forma sistemática por toda la superficie arbolada, y en ellas se miden los diámetros de los árboles a la altura de 1,30m (denominados *diámetros normales*) de todos los pies de árbol incluidos dentro de la parcela, midiéndose además en una submuestra de árboles otras variables más costosas de obtener, como la altura, diámetros de copa, etc. (Mandallaz y Ye, 1999). A partir de esta submuestra de árboles se utilizan ecuaciones para el cálculo del volumen de madera u otras variables de interés en función del diámetro normal, dichas ecuaciones se aplican al resto de árboles medidos en la parcela para obtener los valores por parcela, a partir de los cuales mediante técnicas de estimación tradicionales basadas en el diseño de muestreo aleatorio o mediante técnicas geoestadísticas desarrolladas en las últimas décadas se estima la densidad, normalmente mediante el área basimétrica (superficie de la sección de los troncos de los árboles a la altura de 1,30m referida a una hectárea) o la producción (Montes y col. 2005). A partir de inventarios sucesivos, que en los montes ordenados suelen realizarse cada 10 años, se conoce la evolución de la masa forestal en el tiempo y se planifican los tratamientos selvícolas a realizar.

Hasta la fecha, todas estas medidas se han venido realizando generalmente *in situ*, de forma que sobre las parcelas de muestreo se toman en campo las medidas pertinentes mediante los dispositivos apropiados, lo que hace que el inventario sea una tarea muy costosa, por lo que en la actualidad se están buscando técnicas, como el LIDAR o la teledetección, que consuman menos recursos y permitan tener la información en intervalos de tiempo más cortos. A partir de la utilización del dispositivo MU-200501738 por parte del equipo de investigación forestal del INIA, algunas de tales medidas se vienen realizando de forma semiautomática. Para ello, cada una de las parcelas de muestreo se encuentra perfectamente identificada y localizada mediante posicionamiento espacial a base de marcas sobre el terreno (*landmarks*) y ayuda de GPS (*Global Position System*). Cada vez que se quiere realizar una medición sobre dicha parcela, con el fin de



realizar cualquier análisis comparativo, respecto de mediciones previas, se fija siempre el dispositivo en la misma localización espacial. Además, su posicionamiento relativo debe ser siempre el mismo, de este modo, una imagen, digamos la izquierda del par estereoscópico, siempre se obtiene con la misma orientación espacial del dispositivo, con la cámara situada en el extremo orientado hacia el Sur geográfico de la guía de desplazamiento; la segunda imagen o imagen derecha del par se obtiene con la cámara situada en el extremo Norte de la guía, tras su desplazamiento de Sur a Norte. Nada impide que la orientación sea en relación a cualquiera de las coordenadas geográficas; eso sí, siempre que se quiera realizar una medida comparativa, en diferentes instantes de tiempo, debe ser siempre la misma. El eje óptico de la lente se mantiene paralelo antes y después del desplazamiento de la cámara, por lo que las dos imágenes capturadas constituyen lo que se denomina un *par estereoscópico*, de manera que conociendo el desplazamiento relativo entre dos puntos del par de imágenes estereoscópicas, que representan la misma entidad física tridimensional, por ejemplo un determinado punto a cierta altura en el tronco del árbol, se puede determinar la posición relativa de dicho punto en el espacio 3D respecto del dispositivo de medida.

Actualmente, el equipo de investigación del INIA determina dichos desplazamientos relativos entre puntos homólogos del par estereoscópico mediante un programa informático basado en el trazado de líneas y puntos sobre la pantalla del computador. Esta tarea requiere, naturalmente, la presencia de un operador, encargado de determinar qué punto de una imagen representa la misma entidad física en la otra imagen del par estereoscópico. Una vez establecidos estos desplazamientos relativos, conocidos técnicamente como *disparidades*, es posible determinar las deseadas distancias desde el dispositivo a las entidades del mundo 3D que han originado los desplazamientos en las imágenes. Por el mero hecho de calcular las disparidades mediante la intervención del usuario, se dice al respecto que las medidas se realizan de forma semiautomática. Desde el punto de vista del trabajo de investigación aquí presentado, el objetivo se centra en la identificación de las texturas correspondientes a los troncos, hasta donde sea posible, de forma que en un futuro las mencionadas disparidades se puedan establecer de forma automática. Y si las texturas de interés son los troncos, las texturas pertenecientes al suelo, cielo u hojas de los árboles son texturas carentes de interés. Ambas están presentes

en las imágenes, por lo que el objetivo real consiste en identificar ambas unas por su interés y otras para su exclusión esa falta de interés.

## 1.2 Identificación del problema

En aras del reconocimiento de las texturas anteriormente mencionadas es necesario definir métodos de clasificación apropiados con tal propósito, existen dos problemas claramente identificados en este tipo de tipo de clasificadores. A saber, uno es el que se deriva de la naturaleza supervisada misma de los clasificadores, ya que requiere la presencia de un supervisor, generalmente el usuario. El segundo, la mejora de los resultados obtenidos por dichas estrategias.

Para resolver la problemática presente se ha intentado la búsqueda de soluciones con éxito dispar. Para evitar la presencia del supervisor y conseguir la automatización de las estrategias, se han propuesto métodos que determinan la mejor partición o clasificación de los datos en base a la realización de pruebas hasta encontrar resultados satisfactorios (Duda y col., 2001). En cuanto a la mejora de los resultados obtenidos, la tendencia actual, según se desprende de la revisión bibliográfica realizada en el capítulo dos, se inclina por la combinación de clasificadores bajo la hipótesis de que la suma de esfuerzos mejora los resultados individuales. En esta línea, la comunidad científica viene realizando un importante esfuerzo desde hace más de una década (Kuncheva, 2004; Kittler y col., 1998).

La clasificación de texturas en imágenes naturales de exterior no es ajena a la problemática general planteada. A partir de la imagen de la figura 1.3(a) se observa que la característica más relevante es que presenta una alta variabilidad de los niveles de intensidad en las tres componentes espectrales  $R$ ,  $G$ ,  $B$  por el hecho de haber sido captada en un día soleado. En cambio la imagen de la figura 1.3(b) se caracteriza por haberse obtenido bajo condiciones de escasa iluminación en un día nublado o al atardecer o anochecer en un día soleado. En cualquier caso, aún teniendo como meta el problema de la clasificación, queda patente que dada la problemática presentada por dichas imágenes, las técnicas de clasificación deben diferir necesariamente cuando se trata de clasificar las texturas presentes en uno u otro tipo de imágenes.

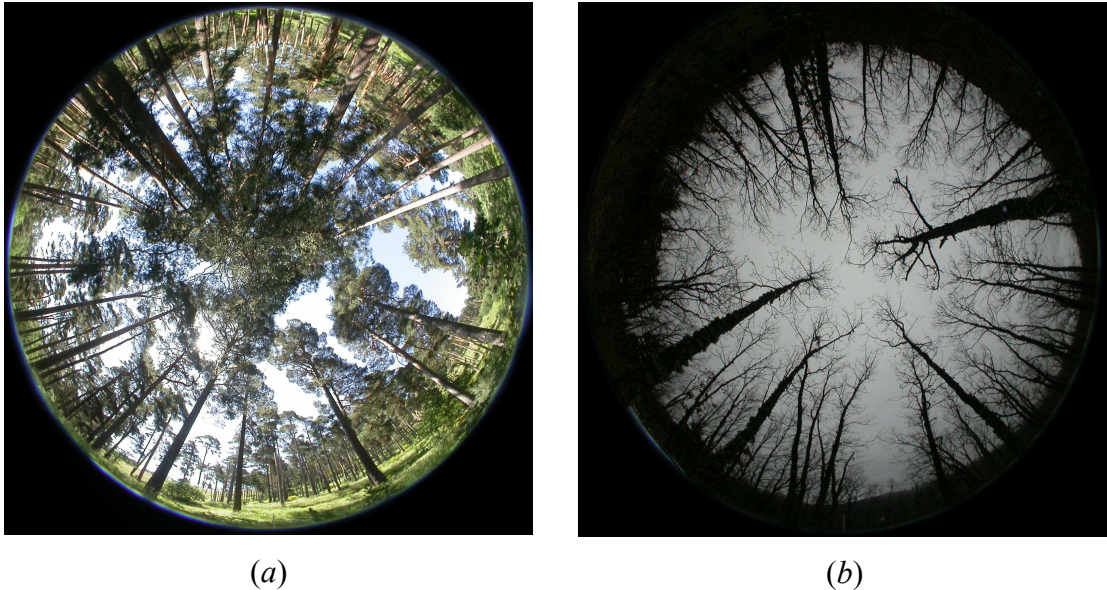


Figura 1.3 Imágenes originales hemisféricas captadas: (a) con alta variabilidad de la intensidad; (b) con escasa iluminación.

Conviene recordar que los troncos de los árboles son los componentes básicos desde el punto de vista del análisis forestal, ya que por un lado contienen el mayor volumen de madera del árbol y por otro identifican la posición del árbol mismo en el área bajo análisis, a partir de los cuales se pueden estimar variables tales como la densidad de árboles, alturas de los mismos u otras ya mencionadas.

Centrando la atención en primer lugar sobre las imágenes del primer tipo, figura 1.3(a), Resulta fácil apreciar, a partir de esta imagen representativa del conjunto de imágenes disponibles, cómo los rayos incidentes del sol producen una importante variabilidad en la iluminación en una parte de la imagen (zona superior derecha), de tal forma que los troncos en esta zona presentan tonalidades claras por su exposición directa al sol. En cambio, el resto de troncos en la imagen presentan tonalidades oscuras por no encontrarse expuestos directamente al sol, y por tanto en la sombra.

Debido a las diferentes características espectrales mostradas y a la problemática exhibida por este tipo de troncos según su ubicación en el espacio y por tanto en la imagen, su identificación mediante este procedimiento no resulta apropiada, entre otras razones porque las imágenes no presentan patrones de textura homogéneos que faciliten su reconocimiento. Obsérvese en la imagen 1.3(a) las diferentes texturas exhibidas por los

troncos aún a pesar de pertenecer a la misma especie de árbol. Por ello, es necesario ahondar más en el estudio y análisis de la problemática presente en estas imágenes.

Tras un análisis más en profundidad sobre este tipo de imágenes, además de las texturas propias de los troncos, se pueden distinguir básicamente otros cuatro tipos de texturas, a saber:

- Vegetación en el suelo (arbustos, hierba, etc.).
- Cielo entre los huecos de las copas de los árboles
- Hojas de los árboles.
- Suelo desnudo (sin vegetación).

Desde el punto de vista del procesamiento de las imágenes y considerando que el interés se centra en la identificación de los troncos, es evidente que estos cuatro tipos de texturas no constituyen el objetivo del proceso posterior de la visión estereoscópica, por lo que nuestro interés sobre ellas se centra precisamente en su identificación para su exclusión. Por este motivo, este trabajo de investigación se planea como objetivo establecer los métodos necesarios para la identificación de estas texturas, de forma que no intervengan en los procesos posteriores, facilitando de este modo lo que será el proceso de correspondencia para los troncos de los árboles.

Hasta donde se ha podido investigar, no existe gran cantidad de trabajos previos específicos que permitan establecer un punto de partida concreto. Uno de ellos es Herrera (2010) donde se proponen una serie de estrategias, a nivel de píxeles y de regiones, para establecer la correspondencia entre los troncos presentes en pares de imágenes estereoscópicas, caracterizadas por su proyección hemisférica. En este trabajo, se propone un estudio relativo a la identificación de las regiones pertenecientes a los troncos, bajo un enfoque basado en píxel y otro basado en regiones. En concreto, en el enfoque basado en píxeles, se proponen una serie de métodos de clasificación con el objeto de identificar una serie de texturas carentes de interés para el proceso de correspondencia posterior. Los métodos allí propuestos sirven de partida para la investigación llevada a cabo y reflejada en esta memoria.

En Guijarro (2008, 2009) se proponen una serie de estrategias relativas a la identificación y clasificación de texturas en entornos naturales. Las imágenes sobre las que se han aplicado son imágenes procedentes de sensores aerotransportados, figura 1.1(a). Las imágenes forestales utilizadas en el presente trabajo y aquellas tienen en común que son imágenes de entornos naturales de exterior y por tanto no estructurados, por lo que los resultados allí alcanzados han servido también de base para la investigación aquí presentada. En concreto, la propuesta de combinación de clasificadores proviene de esas investigaciones.

### 1.3 Motivación

El trabajo de investigación desarrollado tiene su origen en el interés suscitado por el Centro de Investigación Forestal (CIFOR) del Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) para automatizar los procesos de medición mediante el dispositivo con número de patente MU-200501738, mencionado previamente.

De la necesidad suscitada por parte del CIFOR para explotar la potencialidad de dicho dispositivo de medida de cara a la automatización de los procesos para determinar medidas útiles en inventarios forestales, surgió un proyecto de colaboración entre el grupo ISCAR (2006) y el CIFOR titulado *Análisis estereoscópico y reconocimiento de texturas en imágenes forestales obtenidas mediante el dispositivo de medición MU-200501738* al amparo del artículo 83 de la LOU con Ref. INIA 445/2007. Se trata de un sistema de visión estereoscópica con el que se pretende determinar las distancias a las que se encuentran los árboles y a partir de ahí su altura y volumen de madera, entre otras medidas. Para establecer las correspondencias estereoscópicas es necesario identificar las texturas existentes en la imagen con el fin de clasificarlas, bien como troncos o ramas de los árboles, que constituyen los objetos de interés, o bien como otros tipos de texturas tales como suelo, cielo, hojas de árbol, matorrales o hierbas, que deben ser excluidas por carecer de interés e incluso perjudicar el proceso de correspondencia. Sobre la base de este proyecto se mantiene una estrecha colaboración entre el grupo ISCAR y el CIFOR, con el propósito de perfeccionar y ampliar el campo de investigación en el mayor número de entornos forestales posible.

Por otro lado, el grupo ISCAR tiene abierta una línea de investigación en tareas relativas a la clasificación de texturas naturales, como consecuencia del desarrollo de varios proyectos donde dichas tareas se incluyen como fundamentales. La relación entre esas tareas y la investigación desarrollada en este trabajo surge como consecuencia de la colaboración en las mismas en virtud de los correspondientes contratos suscritos entre la Universidad Complutense a través de su Oficina de Transferencia Tecnológica y las respectivas empresas o entidades. También por las necesidades del grupo en las tareas de investigación en el caso de los proyectos del Plan Nacional de I+D en los que participa. A continuación se detallan las líneas generales de tales proyectos, incidiendo en los objetivos relacionados con la identificación de las texturas naturales:

1. Durante el trienio 2009-2011 se colabora en el proyecto del Plan Nacional de I+D 2008 titulado *PERcepción vIsual, toma de deCIsión y Actuación en sistemas artificiales complejos: control selectivo de malas hierbas en cultivos agrícolas (PERICIA)*, ref. AGL2008-04670-C03-02/AGR, con el Instituto de Automática Industrial del CSIC. En este proyecto la identificación de las texturas existentes en las imágenes (cultivo, malas hierbas, suelo) es una tarea primordial para localizar espacialmente los rodales de malas hierbas en campos de cultivo de cereal y maíz con el fin de proceder a su pulverización selectiva. El objetivo final del proyecto consiste en minimizar el efecto contaminante medioambiental por el tratamiento con herbicidas de los campos de cultivo, así como disminuir los costes de producción al necesitar una menor cantidad de producto.
2. Con la empresa TCP Sistemas e Ingeniería se participa en dos proyectos titulados respectivamente *AUTOROVER: estudio de autonomía basada en imágenes para “rover” de exploración planetaria* y *Visión estereoscópica para Auto-rover: estudio de autonomía basada en imágenes*. El primero procede de la participación de la citada empresa en la convocatoria pública según la orden 2259/2007 (B.O.C.M. 272 de 15/11/2007) para el fomento de la innovación en el sector aeroespacial de la Comunidad de Madrid con financiación del Fondo Europeo Regional (ref. 04-AEC0800-000035/2008). El segundo corresponde a una ampliación del anterior dentro del programa Nacional de Proyectos de Investigación en el Sector Aeroespacial según la orden PRE/998/2008 (BOE 11/04/2008) del Ministerio de la Presidencia (ref. SAE-20081093). Ambos

proyectos tienen como fundamento la navegación autónoma de robots en la superficie de Marte e incluyen como uno de los objetivos clave, el análisis de las texturas observadas a través del sistema de visión en dicha superficie. El fin último consiste en discernir entre diferentes tipos de estructuras con el objetivo de determinar la navegabilidad de dichas superficies para los vehículos autónomos.

## **1.4 Objetivos de la investigación**

A la vista de las consideraciones expresadas en los apartados precedentes, se plantean los siguientes objetivos de investigación:

- 1) Aprender a manejar referencias bibliográficas, así como la forma de abordar las investigaciones.
- 2) Identificar métodos de clasificación de texturas existentes en la literatura.
- 3) Determinar los métodos más relevantes en el ámbito de la investigación.
- 4) Analizar las ventajas e inconvenientes de cada uno de ellos así como su problemática.
- 5) Identificar una posible mejora de los métodos existentes o dar una solución novedosa a la problemática de la clasificación de las texturas naturales.
- 6) Identificar líneas de investigación futuras.

## **1.5 Aportaciones a la investigación**

A partir de los objetivos propuestos y los aspectos destacados previamente, se trata de dar solución a los problemas de clasificación planteados en entornos naturales forestales y proponer a la comunidad científica una serie de estrategias de solución que puedan extenderse a otros problemas de naturaleza similar.

La principal contribución de este trabajo de investigación es la combinación de una serie de clasificadores mediante el método de la Toma de Decisiones Multicriterio bajo el paradigma de la lógica Fuzzy. Los clasificadores son los siguientes:

- Un procedimiento de clasificación no supervisada a partir del método de Agrupamiento borroso supervisado.
- Clasificador paramétrico Bayesiano.
- Ventana de Parzen.
- Algoritmo generalizado de Lloyd.
- Mapas Auto-Organizativos.

## 1.6 Organización de la memoria

La presente memoria se presenta estructurada en cinco capítulos:

Capítulo 1. *Introducción*. Se exponen los objetivos que se plantearon en la investigación aquí recogida, en base a los planteamientos previos y que tratan de dar solución a la problemática existente. También se esbozan las estrategias realizadas para abordarlos.

Capítulo 2. *Estado del arte*. Se realiza una revisión del estado del arte en el tema de clasificación de texturas.

Capítulo 3. *Estrategia combinada de clasificación*. Se propone la estrategia de solución a la problemática planteada, junto con una breve revisión de los métodos empleados para llevarla a cabo.

Capítulo 4. *Análisis de resultados*. Se diseña una estrategia de pruebas con el objetivo de verificar la eficacia del método propuesto cuando se compara con otras estrategias existentes.

Capítulo 5. *Conclusiones y Trabajo futuro*. Se extraen las conclusiones pertinentes y se exponen las líneas de investigación futuras.



# Capítulo 2

## Estado del arte

### 2.1 Visión general

En el capítulo uno se ha establecido como objetivo principal la automatización del proceso para la obtención en el futuro de las distancias a puntos significativos de la escena 3D desde el dispositivo de medición MU-200501738 con el fin de poder estimar una serie de parámetros tendentes al análisis y la realización de inventarios forestales.

Como también se ha mencionado en el mismo capítulo, dicho dispositivo es exactamente un sistema estereoscópico basado en imágenes, por tanto, el proceso de automatización pasa por el procesamiento de dichas imágenes bajo el ámbito de cualquier proceso basado en visión estereoscópica pasiva. Aunque las imágenes poseen las características propias de los sistemas ópticos basados en lentes de ojo de pez, el proceso estereoscópico consta de los pasos propios de cualquier sistema estereoscópico convencional basado en proyección de perspectiva, sistemas éstos que han sido por tradición los estudiados ampliamente en la literatura, en contraposición a los basados en lentes de ojo de pez sobre los que la base documental, hasta donde se ha podido investigar, resulta ser más bien escasa, razón por la cual el análisis bibliográfico realizado en este capítulo se basa en la información relativa a los sistemas estereoscópicos convencionales.

Por tanto, independientemente del sistema utilizado, convencional o basado en ojos de pez, siguiendo la terminología establecida en Barnard y Fishler (1982) o Cochran y Medioni (1992), el proceso de la visión estereoscópica consta de los siguientes pasos:

- a)* adquisición de imágenes,
- b)* geometría del sistema,
- c)* extracción de características,
- d)* correspondencia de características.

e) obtención de la distancia.

La adquisición de imágenes es el proceso por el cual se obtienen las dos imágenes del par estereoscópico, en nuestro caso las imágenes con la proyección propia de los sistemas basados en lentes de ojo de pez, tal y como se mostraron en el capítulo uno, figuras 1.2 y 1.3.

La geometría del sistema determina las relaciones físico-geométricas del dispositivo, que definen ciertos parámetros que afectan al proceso global de la visión estereoscópica.

La extracción de características se refiere al proceso por el cual se identifican en las imágenes los elementos propios de las mismas, que se emparejarán entre ambas imágenes y que representan las proyecciones de los elementos reales en la escena 3D. De manera que un elemento, por ejemplo un árbol en la escena, genera sendas proyecciones en las dos imágenes del par estereoscópico. Dichas características pueden ser de diversos tipos, desde los píxeles que forman la imagen en general, pasando por píxeles específicos como los que delimitan puntos de borde hasta agrupaciones de píxeles formando regiones según alguna propiedad común de los mismos. Desde el punto de vista de la investigación desarrollada en este trabajo y por las razones expuestas en el capítulo uno, las características utilizadas serán los píxeles. Por otro lado, aparte de identificar los píxeles pertenecientes a los troncos de los árboles, también se incluyen como características aquellos píxeles que no forman parte de los troncos sino de aquellas otras texturas cuyo objetivo, dentro del proceso, es su exclusión, mediante las técnicas de clasificación previstas, del proceso de emparejamiento y favorecer así el emparejamiento de los troncos. Al proceso completo de extracción de características, extracción de propiedades de las características y clasificación de texturas para su exclusión se le identifica, desde el punto de vista de esta investigación, como proceso de *segmentación*.

La correspondencia de características es el proceso por el cual dada una característica en la escena 3D se llega a determinar cuál es su proyección en sendas imágenes del par estereoscópico. Es decir, se trata de emparejar características homólogas de una imagen con las de la otra imagen del par estereoscópico, que representan la misma entidad física en la escena 3D. Para lograr los emparejamientos, se aplican una serie de restricciones conocidas en la literatura especializada como *restricciones estereoscópicas*. Este proceso se conoce como *correspondencia* propiamente dicho. Una vez establecida la

correspondencia entre dos características, resulta fácil determinar las ubicaciones en las imágenes de ambas características y a partir ahí determinar las separaciones relativas entre ellas, esto es, la *disparidad*, que en el caso de las imágenes analizadas se traduce en diferencias entre los dos ángulos que determinan las posiciones de las características.

Finalmente, una vez establecidas las correspondencias entre características homólogas y por tanto sus disparidades, resulta relativamente fácil determinar las distancias a las que se encuentran en la escena 3D las entidades representadas por dichas características sin más que recurrir a ciertos parámetros conocidos y proporcionados por la geometría del sistema estereoscópico.

De lo expuesto anteriormente, el proceso de segmentación es uno de los procesos clave del proceso global de correspondencia estereoscópica en general y en particular para el tipo de imágenes utilizadas en este trabajo de investigación. Por este motivo, la investigación que se presenta se centra exclusivamente en proporcionar una serie de estrategias y procedimientos encaminados a dar solución al problema de la segmentación para los tipos de imágenes descritas en el capítulo uno. En la figura 2.1 se presenta un diagrama esquemático del proceso global de visión estereoscópica, donde como puede observarse bajo el epígrafe de segmentación de imágenes, se agrupan los procesos de extracción de características y de los atributos de dichas características, cuya salida constituye la entrada al proceso de correspondencia.

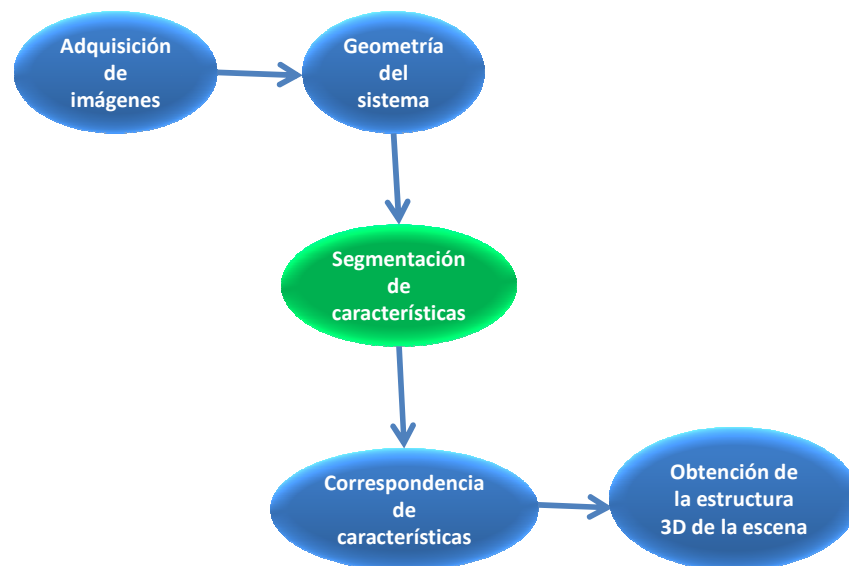


Figura 2.1 Esquema del proceso de visión estereoscópica.

En resumen, el trabajo de investigación que se presenta en este trabajo aborda de forma directa y específica el proceso de segmentación. En él, se encuadran aquellas técnicas encaminadas a la identificación de características a nivel de píxel donde la segmentación constituye un paso clave y esencial. En este caso su objetivo se centra en la identificación de los píxeles pertenecientes a una serie de texturas para su exclusión, que facilitarán sin duda en el futuro el proceso de correspondencia relativo a los píxeles de los troncos. Este hecho se pone de manifiesto en los resultados mostrados en el trabajo de Herrera (2010). Como se refleja en dicho trabajo, el problema de la segmentación en este tipo de imágenes permanece abierto a la investigación, por este motivo resulta absolutamente imprescindible realizar un estudio bibliográfico en profundidad sobre métodos de identificación y clasificación de texturas en entornos naturales, ya que hasta donde se ha podido investigar no existen técnicas específicas que den solución a dicha problemática. También se debe justificar dicho estudio bibliográfico sobre la base de la búsqueda de mejoras en las propuestas formuladas en este trabajo.

En este capítulo se recopila información sobre la utilización de combinaciones de clasificadores aplicados a imágenes digitales; en ocasiones, a este tipo de clasificadores se les denomina *híbridos* en la literatura especializada. En cualquier caso, de forma general, un clasificador combinado tiene su fundamento en la participación de otros clasificadores individuales, que constituyen las unidades básicas de la combinación. Se analizan los aspectos relacionados con la combinación de clasificadores desde las perspectivas más comúnmente utilizadas en la literatura, orientándose hacia la clasificación de texturas en imágenes naturales de exterior, que constituyen el objetivo de esta investigación. Los aspectos tratados se estructuran como sigue. En la sección 2.2 se revisan distintos métodos genéricos encaminados al reconocimiento de texturas. La sección 2.3 realiza un estudio relativo a la conveniencia de utilizar métodos de clasificación combinados. En la sección 2.4 se estudian métodos cuyo fundamento son los píxeles como unidades básicas de las imágenes o agrupaciones de los mismos según ciertas distribuciones regulares. Por tanto, en esta categoría se incluyen tanto las propiedades espectrales (color), como espaciales (distribuciones de texturas) o ambas bajo un enfoque combinado. Posteriormente, en la sección 2.5 se analizan en concreto métodos cuyo fundamento son las relaciones de vecindad existentes entre un píxel y los que le rodean. Aquí se

contemplan las agrupaciones irregulares espaciales de píxeles, que son realmente tal y como aparecen en las imágenes naturales objeto de este trabajo de investigación.

Finalmente, como consecuencia de la revisión previa, en la sección 2.6 se incluye un resumen sobre aquellas técnicas más relevantes desde el punto de vista del trabajo que se presenta, constituyendo por tanto, el origen de la investigación desarrollada y que se detalla en capítulos sucesivos.

La mayoría de los métodos existentes en la literatura utilizan más de una estrategia o método de clasificación con diferentes características. Por tanto, no resulta fácil ubicar un determinado trabajo de investigación exclusivamente en un epígrafe, razón por la cual algunos de los métodos expuestos aparecen citados en varias secciones; no obstante se ha procurado encuadrar cada uno de ellos según la aportación más relevante de acuerdo con el contenido de la sección.

## **2.2 Métodos de clasificación de texturas en imágenes**

Durante las últimas tres décadas e incluso cuatro, se han propuesto en la literatura diversos métodos encaminados a la segmentación de las imágenes según sus texturas, ver por ejemplo Xia y col. (2007) y referencias asociadas. En Wilkinson (2005) se realiza una exhaustiva revisión de métodos de clasificación durante los quince años previos a la publicación de dicho trabajo aplicados a las imágenes de satélite.

La mayoría de ellos pueden considerarse como métodos con dos procesos sucesivos asociados: *extracción de características* y *segmentación basada en las características*. El término sucesivos hace referencia al hecho de que se realizan en ese orden durante la clasificación.

La extracción de características tiende a encontrar una medida apropiada para caracterizar la homogeneidad de cada región dentro de la imagen, estableciendo una asociación entre la región y una determinada textura (Valdovinos y col. 2005, Puig y García, 2006 o Hanmandlu y col. 2004). Algunas de tales características se analizan posteriormente en este mismo capítulo en la sección 2.4.

La segmentación basada en las características consiste en el proceso de asignar a cada píxel de la imagen una etiqueta que lo identifique como perteneciente a una determinada

clase. Además, se han diseñado y utilizado una amplia gama de algoritmos tanto de naturaleza supervisada como no supervisada. A continuación se realiza un esbozo de los más relevantes, si bien algunos de ellos serán estudiados con mayor profundidad en el capítulo tres, dado que constituyen la base de la propuesta de combinación realizada en este trabajo.

### *1. Técnicas borrosas o fuzzy*

Asumen el hecho de que los colores tienden a agruparse en la imagen formando clases que determinan la distinta naturaleza de las texturas. Dada la naturaleza *fuzzy*, en cada una de las clases, se admiten píxeles con cierta ambigüedad, en el sentido de que un mismo píxel podría pertenecer tanto a una textura como a otra según sus características. Este hecho se manifiesta más claramente en las zonas fronterizas de las texturas, donde se producen transiciones graduales en los valores de los píxeles de una zona a otra de la frontera. A la categoría *fuzzy* pertenece uno de los métodos más clásicos en clasificación como es el conocido Agrupamiento borroso o en inglés “*Fuzzy Clustering*”. Para la fase de entrenamiento utiliza un método de optimización, que permite conseguir agrupaciones compactas y resolver muchas de tales ambigüedades. En esta línea se encuentran los trabajos de Pajares y Cruz (2002), Pajares y col. (2002) o Lim y Lee (1990) entre otros.

Mohaddamzadeh y Bourbakis (1997) definen las regiones a segmentar como agrupaciones de píxeles que poseen similar color, mientras los cambios de región se manifiestan gradualmente como ya se ha indicado previamente. Para plasmar esta idea definen una función de naturaleza *fuzzy* en el espacio de color *RGB* (Pajares y Cruz, 2007).

Cheng y col. (2001) utilizan el concepto de “*homogram*” para extraer regiones homogéneas en una imagen de color; a su vez definen el concepto de homogeneidad *fuzzy* para extraer regiones de similares características mediante la obtención de una serie de umbrales. Tras la segmentación vuelven a aplicar un criterio de similitud para fusionar regiones, se trata de un proceso similar al método de segmentación de regiones según la técnica de árboles cuaternarios (Pajares y Cruz, 2007).

En el ámbito de las imágenes de satélite los siguientes trabajos utilizan métodos de clasificación basados en la lógica *fuzzy*: Seong y Usery (2001), Zhang y Foody (1998).

## 2. Redes neuronales

Por el carácter no lineal, que permite una mayor separabilidad en el espacio de características y su alta capacidad para desarrollar procesamiento en paralelo, las redes neuronales se han utilizado también en la clasificación de texturas en imágenes en general y en imágenes de satélite en particular (Mas y Flores, 2008).

Dentro de las redes neuronales se encuentra la red neuronal de Hopfield (Haykin, 1994), compuesta por una serie de neuronas o nodos conectados entre sí. Este tipo de red realiza un proceso de optimización basándose en la minimización de una función de energía hasta que la propia red adquiere un grado de estabilidad aceptable. En Huang (1999) se detectan los picos más significativos del histograma y se diseñan tres redes de Hopfield, una para cada banda de color, donde los nodos de cada red son los píxeles de la correspondiente banda espectral y las conexiones establecen similitudes entre ellos. En Campadelli y col. (1997) se diseña un algoritmo con  $M \times N \times S$  neuronas siendo  $M$  y  $N$  las dimensiones de la imagen y  $S$  el número de clases que se obtienen mediante el análisis previo del histograma. En este tipo de redes se determinan las conexiones entre los nodos mediante el cómputo de funciones de consistencia entre los píxeles, basándose en las etiquetas de los vecinos. Para ello, se requiere la realización de un proceso de clasificación y etiquetado previo.

La red retro-propagación ha sido utilizada en clasificación de imágenes naturales, cabe mencionar entre otros los trabajos de Pajares y col. (2001), Atkinson y Tatnall (1997) o Paola y Schowengerdt (1995), donde las componentes espectrales correspondientes a los tres colores básicos de cada píxel constituyen las entradas a la red, mientras que como salidas se establecen las clases a las que pertenecen dichos píxeles según un mecanismo de aprendizaje supervisado.

Los mapas Auto-Organizativos (*Self Organizing feature Maps*, SOM) han sido otra de las herramientas utilizadas en clasificación. Este es el caso de los trabajos desarrollados en Jiji y Ganesan (2008) o Lo y Pei (1999), entre otros. En ellos los centros de las clases se estiman según la correspondiente ley de aprendizaje. Este método se describe exhaustivamente en el capítulo tres, sección 3.2.5, debido a que se trata de uno de los clasificadores que participa en la estrategia combinada propuesta en este trabajo de investigación.

### 3. *Cuantización vectorial ("Learning Vector Quantization")*

Se trata de un procedimiento que puede ser de naturaleza supervisada o no supervisada. En su versión supervisada se determina de antemano el número de clases existentes en la imagen, mientras que en la no supervisada es a través del propio método como se establece el número de clases en función de un umbral fijado previamente. En cualquiera de las dos versiones, se trata de un proceso de aprendizaje, que obtiene los centros de las clases. Algunos trabajos basados en estos métodos, que además utilizan información del color, son los propuestos en Giacinto y col. (2000), Chan y col. (2003) o del Frate y col. (2007). Los trabajos de Lillo y col. (2007), Aiyer y col. (2005) o Pyun y col., (2002) realizan una cuantización vectorial utilizando propiedades de texturas obtenidas mediante la aplicación de la transformada de Fourier (Pajares y col. 2007; Gonzalez y Woods, 2007).

### 4. *Vecinos más próximos*

Este método se conoce en la literatura inglesa como *k-Nearest Neighbours* (*k*-NN). En ellos, dado un conjunto de clases, los nuevos píxeles a clasificar con sus características, se asignan a aquella clase que posea un número de *k* vecinos más próximos. Dicha proximidad se determina mediante alguna medida de similitud, por ejemplo una distancia. En Haapanen y col. (2004) se utiliza esta técnica para la segmentación de zonas boscosas en imágenes aéreas. En Muñoz-Marí y col. (2007) se utiliza este clasificador como referencia para el análisis comparativo de su método. En Singh y col. (2001) se aplica este método para clasificar nueve tipos diferentes de texturas.

### 5. *Estimas de máxima verosimilitud*

Consistentes en la estimación de una función de densidad de probabilidad, que constituye el objeto del aprendizaje. Tomando como referencia la función de densidad de probabilidad estimada, es posible determinar la probabilidad de pertenencia de una muestra dada a las diferentes clases. En esta línea se encuentran los trabajos de Frizelle y Moody (2001), Newsam y col. (2004), Xie y Tsui (2004), Trauwaert y col. (1991) o Benediktsson y col. (1990). Muñoz-Marí y col. (2007) utilizan también este clasificador para comparar los resultados obtenidos con su método.



## 6. *Algoritmos genéticos*

Utilizados como métodos de optimización que aprovechan la información espacial de la vecindad para definir las funciones objetivo, tal es el caso de Tso y Mather (1999).

## 7. *Máquinas de vectores soporte*

Estas técnicas están especialmente diseñadas para aplicaciones bi-clase, si bien se han propuesto ampliaciones para abordar el problema multi-clase. Tal es el caso de los trabajos propuestos en Camps-Valls y Bruzzone (2005), Camps-Valls y col. (2004) o Melgani y Bruzzone (2004). En Ferreiro-Armán y col. (2007) se realiza un estudio sobre el comportamiento en imágenes de satélite de las máquinas de vectores soporte, además del Agrupamiento borroso, los Mapas Auto-organizativos y la red neuronal retro-propagación.

## 8. *Árboles de decisión*

Se presentan como una alternativa a los métodos clásicos de clasificación donde lo que se infieren son reglas de decisión, que permiten realizar las clasificaciones en base al cumplimiento de las premisas aprendidas y reflejadas en las reglas (Tottrup, 2007; Kumar y Majunder, 2001; Hansen y col. 1996). En Debeir y col. (2002) se utiliza el esquema C4.5 para la clasificación de imágenes aéreas.

Un aspecto importante a destacar respecto de los métodos mencionados en los puntos uno a ocho, es el hecho de que se centran exclusivamente en la clasificación de los píxeles sin tener en cuenta las características espaciales de las imágenes. Así, para cada píxel de la imagen se extraen las correspondientes propiedades o características, procediendo seguidamente a su clasificación sin tener en cuenta su posición o relación con otros píxeles vecinos en la imagen. Con el fin de incorporar la información espacial, se han realizado diferentes esfuerzos para considerar esta importante propiedad de las imágenes. En efecto, una imagen se define como una estructura rectangular donde cada píxel se ubica en una determinada localización espacial dentro de la imagen. Después de extraer las características de cada píxel, la imagen puede expresarse como agrupaciones de características, de forma que cada agrupación define una región. Cada región posee características similares, además de una cierta compacidad espacial y probablemente, constituya una parte de un objeto, por ejemplo una textura. De aquí surge la idea de que

durante el proceso de clasificación no sólo se deben tener en cuenta las características individuales de los píxeles sino también su localización e información espacial. En este sentido, se han propuesto diferentes estrategias para incorporar la información espacial en los métodos de clasificación. Posteriormente, en la sección 2.5 de este mismo capítulo, se aborda específicamente este aspecto.

## **2.3 Métodos de clasificación combinados**

Una importante línea de investigación actual para mejorar los resultados de la clasificación obtenidos mediante clasificadores individuales, consiste en combinar éstos. La combinación de clasificadores individuales constituye el objetivo de esta sección.

En primer lugar, se abordan los aspectos relacionados con la combinación desde una perspectiva general, para posteriormente tratar de forma específica la combinación para la segmentación de las texturas en imágenes.

### **2.3.1 Clasificadores combinados de índole general**

Un aspecto importante a destacar, asumido por la comunidad científica en general, es el hecho de que mediante la combinación de clasificadores se obtienen mejores resultados que con los clasificadores utilizados de forma individual, esto es lo que se concluye en los trabajos de Valdovinos y Sánchez (2007), Valdovinos y col. (2005), Kuncheva (2003, 2004), Kumar y col. (2002), Kittler y col. (1996, 1998) o Cao y col. (1995) entre otros muchos. Por otro lado, además existen diferentes estudios relativos al tema de la combinación, en ellos se destacan las ventajas de este planteamiento en el sentido de que la combinación permite resaltar las bondades de los clasificadores al mismo tiempo que se atenúan o desaparecen las desventajas. Según los resultados obtenidos en los trabajos citados, la justificación sobre la conveniencia de utilizar clasificadores queda suficientemente acreditada. Surge, en este momento, la cuestión de cómo llevarla a la práctica. En este sentido, destacan dos importantes estrategias, comúnmente aceptadas por la comunidad científica, para llevar a cabo la combinación de clasificadores, a saber (Valdovinos y col., 2005): *selección* y *fusión*.

En la selección, cada clasificador individual realiza una decisión y se elige un único clasificador como el mejor de acuerdo a algún criterio preestablecido. Por el contrario, en

la fusión, todos los clasificadores participan de una u otra forma según la estrategia de combinación utilizada.

En cualquier caso, existen algunas estrategias donde la combinación se realiza mediante una o más redes neuronales, que reciben como entradas las salidas de los clasificadores individuales con los que realizan el correspondiente proceso de aprendizaje. En esta línea se encuentran los trabajos de Wozniak y Jackowski (2008), Hashem y Schmeiser (1995) o Hansen y Salamon (1990).

### **2.3.2 Combinación de clasificadores en texturas de imágenes**

En el ámbito de la clasificación de texturas naturales en imágenes, se pueden mencionar algunos trabajos de interés que han utilizado combinaciones de algunas de las técnicas anteriormente expuestas. En este sentido se encuentran los trabajos de Lepisto y col. (2003, 2005), quienes proponen un nuevo método de clasificación donde las salidas de los clasificadores individuales forman los vectores patrón para el nuevo proceso de clasificación. Dichos vectores contienen las opiniones de los clasificadores individuales en cada una de sus componentes. De esta manera, la decisión se lleva a cabo comparando estas opiniones con respecto a las almacenadas para las muestras y utilizando un criterio de mínima distancia tal como la distancia de Hamming. Este procedimiento ha sido comparado favorablemente frente a los clasificadores combinados clásicos tales como la regla del máximo, media o el método conocido como votación mayoritaria, propuestos en Kittler y col. (1996, 1998).

Muñoz-Marí y col. (2007) proponen un método combinando los cuatro clasificadores siguientes, que pueden encontrarse descritos en Pajares y Cruz (2007): 1) estimador de máxima verosimilitud, modelado como una función de densidad de probabilidad Gaussiana; 2) funciones de densidad de probabilidad gaussianas mixtas; 3)  $k$ -vecinos más próximos y 4) máquinas de vectores soporte. Estos clasificadores se entrenan de forma independiente para cada una de las clases identificadas. Posteriormente, para la clasificación, se combinan de forma que la decisión se toma en función de cuál de ellos proporciona la mejor opción. Aparentemente, este esquema parece sencillo, si bien la novedad más importante que introduce consiste en el hecho de que ante la posibilidad de que aparezca una muestra que no pertenece a ninguna de las clases utilizadas durante el entrenamiento, esta muestra se rechaza sin ser asignada a ninguna de las clases. El

método se aplica a imágenes de satélite procedentes de dos sensores, Landsat TM (2009) y SAR (2009) (“*Synthetic Aperture Radar*”). En esta misma línea se sitúa el trabajo de Briem y col. (2002) para el mismo tipo de imágenes.

Fauvel y col. (2006) combinan los clasificadores mediante operadores de agregación *fuzzy* para clasificación de áreas urbanas a partir de imágenes del satélite IKONOS (SIC, 2009). Estos operadores son básicamente tres: combinación conjuntiva (mínimo), combinación disyuntiva (máximo) y un tercero que elige primeramente el mínimo y luego el máximo.

En Debeir y col. (2002) se propone un clasificador combinado, que utiliza la técnica conocida como “*bagging*” (Duda y col., 2001) con un método propio basado en la extracción de subconjuntos de características de los píxeles a clasificar. Los resultados obtenidos se comparan con el método  $k$ -NN, siendo  $k$  igual a cinco y también con el método basado en árboles de decisión C4.5. Se aplican a imágenes de satélite procedentes del sensor Landsat TM (2009).

En Kumar y col. (2002) se plantea la clasificación de imágenes proporcionadas por el sensor AVIRIS (2009). En dicho trabajo se utilizan 180 bandas espectrales, de forma que la dimensión del espacio de características para cada píxel se corresponde con dicho número de bandas. Debido a la alta dimensionalidad de los datos, el número de clases es también elevado, produciéndose fuertes solapamientos entre las clases. La propuesta para la solución a esta problemática la proporciona el citado trabajo donde se establece una jerarquía en forma de árbol comenzando por el nodo raíz, que representa el problema original con todas las clases. El siguiente nivel de la jerarquía contiene una primera partición del conjunto de clases en subconjuntos disjuntos, obtenidos según ciertos criterios de separabilidad entre las clases. El proceso continúa hasta que no se admiten más particiones. De este modo se consigue la separación de las clases, de forma que en cada nivel se estima una función de densidad de probabilidad para cada clase por máxima verosimilitud. Con dicha función y la probabilidad a priori se obtiene la probabilidad a posteriori mediante la regla de Bayes, que proporciona a su vez la probabilidad a priori para el siguiente nivel. La propia regla de Bayes, a través del producto de las probabilidades, constituye el método de combinación.

Petrakos y col. (2001) definen, en primer lugar, tres índices para medir los resultados de la clasificación, los cuales se obtienen comparando dos a dos los clasificadores, para ello

calculan los siguientes índices: muestras clasificadas correctamente por ambos clasificadores, muestras clasificadas correctamente por el primer clasificador e incorrectamente por el segundo, muestras clasificadas incorrectamente por el primero y correctamente por el segundo y muestras clasificadas incorrectamente por ambos. Dichos índices se denominan: *proporción de concordancia*, *proporción de concordancia específica* y *estadística de kappa* ( $\kappa$ ). El método se aplica a imágenes de satélite procedentes de dos sensores, Landsat TM (2009) y SAR (2009), para cada imagen y cada clase a identificar se estima una función de densidad de probabilidad por separado. Luego, dada una clase, ambas funciones se combinan bajo un esquema de promediado ponderado mediante un coeficiente que asigna un peso específico a cada una de las dos funciones. Los resultados se comparan dos a dos entre los diferentes métodos individuales y combinados que estudia.

En Jiménez y col. (1999) se propone la combinación en dos niveles de los datos hiperespectrales procedentes de tres sensores: óptico, térmico e infrarrojo. En una primera instancia, los datos de cada sensor se convierten en un espacio de características apropiado, donde son escalados a un rango de valores. En el segundo nivel, se lleva a cabo la decisión mediante técnicas clásicas de combinación de clasificadores tales como votador mayoritario, reglas del máximo, mínimo y media. El procedimiento combinado se aplica a imágenes de satélite en general.

Bruzzone y col. (1999) proponen una combinación de métodos para clasificar pares de imágenes obtenidas en diferentes instantes de tiempo y procedentes de diferentes sensores, por ejemplo un sensor óptico y otro SAR (2009). Ambas imágenes contienen la misma región espacial. Los datos o características de los píxeles a clasificar procedentes de cada uno de los sensores constituyen las entradas para dos redes neuronales del tipo retro-propagación, las cuales, tras el correspondiente entrenamiento, proporcionan sendas salidas. Con estas salidas se estiman a su vez dos funciones de densidad de probabilidad utilizando el algoritmo conocido como Expectación Máxima (Duda y col., 2001). Previamente se determinan las probabilidades *a priori* teniendo en cuenta la frecuencia de aparición de las distintas clases en las imágenes analizadas. La función de densidad de probabilidad conjunta se obtiene considerando la propiedad de independencia entre ambas, resultando ser el producto de las dos funciones. Ante la llegada de una nueva muestra, se obtienen sendas probabilidades *a posteriori*. Las dos probabilidades se

combinan mediante la regla de Bayes (Pajares y Cruz, 2007) para tomar una decisión respecto de la clasificación para la muestra entrante.

En Wilkinson y col. (1995), Kanellopoulos y col. (1993) o Benediktsson y Kanellopoulos (1999) se propone la combinación de redes neuronales con métodos estadísticos basados en máxima verosimilitud bajo la hipótesis de que ambos métodos producen buenos resultados y por tanto, pueden complementarse. Inicialmente, los dos clasificadores se entrenan con el mismo conjunto de datos. Posteriormente, una segunda red neuronal se vuelve a entrenar sólo con aquellos datos en los que los dos clasificadores previos discrepan. Esta segunda red tiene la misión de discriminar píxeles ambiguos. Estos métodos se han probado en imágenes procedentes de diferentes sensores de satélites.

La aportación realizada en este trabajo de investigación pertenece a la categoría de esta sección, es decir, clasificadores combinados para segmentación de texturas en imágenes naturales de exterior.

## **2.4 Características de las imágenes en la clasificación de texturas naturales**

Por *característica* se entiende un elemento significativo de la imagen, pudiendo ir desde una región formando parte de un objeto hasta un píxel, pasando por puntos significativos tales como bordes o incluso agrupaciones de píxeles siguiendo un cierto patrón. En cualquier caso, cada característica se describe por una serie de propiedades que la identifican frente al resto. Las características y sus propiedades, constituyen las muestras para el proceso de clasificación, incluyendo el entrenamiento y la decisión. En la literatura se ha estudiado la conducta de diferentes características en la clasificación de las texturas (Valdovinos y col., 2005, Puig y García, 2006 o Hanmandlu y col., 2004). Existen dos enfoques: 1) métodos basados en las características espectrales de los píxeles (Puig y García, 2006; Rud y col. 2006; Hanmandlu y col. 2004; Yu y col. 2002; Kumar y col., 2001) y 2) métodos basados en regiones (Puig y García, 2006; Randen y Husøy, 1999; Wagner, 1999; Smith y Burns, 1997). Los métodos basados en los píxeles tratan de asignar el píxel a una clase determinada en función de sus propiedades específicas e individuales. Los métodos basados en las regiones tienen en cuenta las distribuciones espaciales de los píxeles dentro de la textura y por tanto su relación con otros píxeles en

la región. En este último caso, cabe todavía distinguir entre texturas que presentan ciertas distribuciones regulares o patrones repetitivos, donde la unión de patrones define la textura global y aquellas otras que carecen de distribuciones de este tipo, tratándose por tanto, de texturas de naturaleza irregular.

En esta sección se analizan métodos basados en ambos enfoques. Así en la subsección 2.4.1 se revisan los métodos basados en los píxeles, bajo el punto de vista de su tratamiento individual. En la 2.4.2 se abordan los métodos que tienen en cuenta la distribución regular de las texturas, posponiendo el estudio de las texturas de naturaleza irregular para la sección 2.5. La intención es tratarlas por separado, dado que constituyen la base fundamental de este trabajo de investigación, puesto que estas texturas irregulares son las que aparecen en las imágenes naturales utilizadas en los experimentos llevados a cabo por el método propuesto en este trabajo de investigación. Finalmente, en la subsección 2.4.3 se abordan las técnicas que utilizan tanto los píxeles de forma individual como formando parte de regiones.

### **2.4.1 Características espectrales a nivel de píxel**

Como bien es sabido, los elementos básicos constituyentes de las imágenes son los píxeles, en este caso, considerados como características fundamentales de las propias imágenes, cuya propiedad más relevante es su naturaleza espectral. Existen técnicas de segmentación específicamente dedicadas a imágenes monocromas, es decir, imágenes de grises. No obstante, lo más frecuente es la utilización del color como atributo (Koschan y Abidi, 2008; Pajares y Cruz, 2007; Du y col., 2004; Cheng y col. 2001).

El tratamiento del color ha recibido una importante atención en las últimas dos décadas, en su calidad de propiedad descriptora de los píxeles en las texturas (Jiji y Ganesan, 2008; Rud y col. 2006; Luo y col., 1998; Comaniciu y Meer, 1997; Pappas, 1992; Chang y col., 1994). Siguiendo esta línea, se han realizado diferentes estudios relacionados con los modelos de color, destacando los trabajos de Broek y Rikxoort (2004), Yu y col. (2002) o Kumar y col. (2001) donde se analizan entre otros, los modelos de color *RGB* (*R*, red; *G*, green; *B*, blue), *HSI* (*H*, hue; *S*, saturation; *I*, intensity), *YIQ*, *YUV* o *LUV*, todos ellos encaminados hacia la clasificación de texturas. En general, destacan como los más eficientes los modelos *RGB* y *HSI*. El modelo *RGB* es muy utilizado en técnicas basadas en gráficos por computador y específicamente en visión por computador. Posee una gran

---

dependencia de las características del computador donde se visualiza o representa el gráfico o la imagen. Por otro lado, el modelo *HSI* está más bien relacionado con el concepto de percepción humana según se indica en Tkalcic (2003). En la literatura se han considerado otras representaciones de color tales como  $L^*a^*b^*$  o  $L^*u^*v^*$  debido también a su alto contenido perceptual (Yang y col., 2008).

En Trias-Sanz y col. (2008) se hace un estudio sobre diferentes modelos de color para la segmentación de imágenes de satélite correspondientes a entornos rurales, llegando a la conclusión de que en algunos casos resulta más conveniente utilizar ciertos modelos de color transformados, además del modelo *RGB*, que también es empleado.

Un análisis comparativo, con un elevado nivel de profundidad, sobre distintas representaciones del color se puede encontrar en Cheng y col. (2001), incluyendo las características monocromas. En dicho trabajo se concluyen una serie de ventajas e inconvenientes de los diferentes modelos de color.

El análisis de diferentes trabajos, en segmentación de texturas en imágenes, lleva a la conclusión de que los modelos *RGB* y *HSI* compiten ambos por ser utilizados en dicha tarea. Aunque en algunos casos puntuales con el modelo *RGB* se pueden obtener peores resultados que con el *HSI*, bien es cierto que aquél, en general, se presenta como un modelo satisfactoriamente aceptable, incluso a pesar de que sus componentes poseen una cierta dependencia lineal entre ellas (Comaniciu y Meer, 2002).

En algunos trabajos muy preliminares, como el de Robinson (1977), se llegó a la conclusión de que el modelo *HSI* era el más apropiado para la extracción de bordes en imágenes en color, si bien no es éste el problema que nos ocupa en relación a la clasificación de texturas abordado en este trabajo.

Dentro del análisis del color existen otras aproximaciones relativas a la percepción visual, este es el caso de los trabajos propuestos por Vázquez y col. (2007) o Tkalcic y Tasic (2003). En ellos, a partir del histograma se identifican pequeñas variaciones del mismo con el fin de suprimirlas, consiguiendo así un cierto suavizado, y con ello la homogenización de las regiones de texturas.



Montiel y col. (2005) proponen utilizar la información proporcionada por las distribuciones de los niveles espectrales a través del análisis de los correspondientes histogramas de las imágenes, que contienen las texturas.

Wangenheim y col. (2007), siguiendo la metodología basada en la división y fusión de regiones, aplican medidas de similitud para separar las clases que son diferentes y fusionar las que son similares. Las medidas de similitud se establecen en relación a las componentes de color.

### 2.4.2 Características con patrones espaciales

Este tipo de características tienen la particularidad de presentar agrupaciones de píxeles con ciertas distribuciones regulares y espaciales de patrones. Existen diversas bases de imágenes con este tipo de texturas, tal es el caso del álbum de texturas de Brodatz (1966), disponible en numerosos sitios Web actualizada o la base de datos VisTex por citar sólo algunas. En la Figura 2.2 se muestran varios ejemplos de tales texturas.



Figura 2.2 Diferentes ejemplos de texturas procedentes de la base de imágenes VisTex

Para el análisis de dichas texturas se utilizan diversos operadores de vecindad y filtros cuya descripción se puede encontrar en Pajares y Cruz, (2007), Gonzalez y Woods, (2007), Escalera (2001), Wagner (1999), Randen y Husoy (1999) o Smith y Burns (1997) entre otros. La idea consiste en situar una máscara alrededor de un determinado píxel, gracias a la cual se obtiene información sobre la distribución de los píxeles de la textura en la vecindad del píxel dado.

A continuación se ofrece una relación de diferentes métodos utilizados para la descripción de este tipo de texturas:

- Matrices de co-ocurrencia, con una serie de propiedades basadas en el histograma de las matrices, tales como: varianza, entropía o momentos de segundo orden entre otros (Chamorro-Martínez y col., 2007; Debeir y col., 2002).
- Filtros de Gabor (Wan y col., 2007; Palm y Lehmann, 2002; Dunn y Higgins, 1995; Porat y Zeevi, 1989).
- Máscaras de Laws (Gonzalez y Woods, 2007).
- Descomposición mediante la transformada de wavelets (Wang y Boesch, 2007, Pi y col. 2006; Permuter y col., 2006; Arivazhagan y Ganesan, 2003; Chang y Kuo, 1993; Unser, 1995).
- Descomposición en coeficientes de energía (Chen y col. 2005; Manjunath y Ma, 1996).
- Dominio de la frecuencia mediante análisis de Fourier (Lillo y col., 2007; Aiyer y col., 2005; Pyun y col., 2002; Clark y col., 1999).
- Transformada Discreta del Coseno (Li y col., 2006; Permuter y col., 2006).

En algunos casos se fusionan diversas estrategias, como ocurre en el método combinado descrito en Zhang y Wang (2004), que propone una estrategia multi-resolución mediante la descomposición de la imagen original en bandas de frecuencia a través de la transformada de wavelets, llegando a un cierto nivel de descomposición que es donde se utilizan descriptores basados en las matrices de co-ocurrencia (Pajares y Cruz, 2007). En Chen y col. (2005) también se utiliza una estrategia multi-resolución para la descomposición de la imagen en diferentes bandas de frecuencia.

En Trias-Sanz y col. (2008) se han analizado diferentes tipos de descriptores de textura, tales como filtros de Gabor, patrones locales binarios, estimadores de Fourier o intensidad media, incluyendo algunos otros de naturaleza estructural. Las imágenes utilizadas son de satélite con predominio de zonas rurales. En este trabajo se concluye que el comportamiento de dichos descriptores no resulta apropiado para este tipo de imágenes.

Sobre la figura 2.3 pueden apreciarse alguna de las texturas del mismo tipo anteriormente mencionadas en relación a la figura 2.2, destacando el caso de los troncos, que constituyen un objetivo de interés de primera magnitud en nuestra investigación, si bien sobre este aspecto caben las siguientes matizaciones:

1. En la imagen de la figura 2.3 no aparecen patrones de textura repetitivos como en la figura 2.2, que permitan aplicar los descriptores que se usan para el tratamiento de esta última.
2. No existe una única textura, puesto que se distinguen claramente diversas zonas, tales como el cielo, hierba en el suelo, hojas en los árboles y los propios troncos de los árboles e incluso sobre estos últimos aparecen texturas con una importante variabilidad en la intensidad en función de la incidencia o no de los rayos del sol sobre ellos, a pesar de ser texturas de la misma naturaleza, es decir troncos de pinos.
3. Existe una clara distinción de los colores asociados a las bandas espectrales, lo que sin duda facilitará la identificación de determinadas texturas.



Figura 2.3 Ejemplo de imagen hemisférica con diferentes texturas naturales

Diversos experimentos llevados a cabo sobre este tipo de imágenes, utilizando los descriptores mencionados anteriormente, no han producido los resultados que cabía esperar. Ello significa que las texturas presentes en nuestras imágenes no exhiben patrones de textura repetitivos, razón por la que se han descartado para su uso en el

proceso de segmentación para capturar propiedades de los píxeles mediante dichos descriptores. Sin embargo, conectando con lo relativo a las componentes espectrales, el modelo de color *RGB* ha ofrecido resultados ciertamente satisfactorios, incluso el modelo *HSI* parece vislumbrarse como prometedor en futuras investigaciones tal y como se pone de manifiesto en el trabajo de Herrera (2010).

En cualquier caso, el tratamiento de la información espacial, según los patrones de textura vistos aquí, no debe confundirse con los métodos descritos en la sección 2.5, donde las distribuciones espaciales carecen de tales patrones, como se verá allí.

### **2.4.3 Combinación de características espectrales y espaciales**

Las técnicas combinadas surgen bajo la idea de que la unión de esfuerzos resulta más productiva que los aportes individuales. En este caso, se trata de aprovechar la información conjunta proporcionada tanto por las características espectrales como la procedente de los patrones espaciales introducidos en la sección previa. Una herramienta muy utilizada para llevar a cabo esta fusión de información es la transformada de wavelets.

En la línea expuesta previamente se sitúa el trabajo de Chen y col. (2005). En él se utilizan dos tipos de características. Unas describen la composición local del color a nivel de píxel y las otras las características espaciales de textura en la imagen de intensidad. Cada tipo de características se procesa por separado, para posteriormente proceder a su combinación. Las características de textura se obtienen descomponiendo la imagen en diferentes bandas de frecuencia, de forma que en cada nivel de la descomposición se utilizan aspectos relacionados con la frecuencia espacial.

En Wang y Boesch (2007) se combinan, tanto descriptores de color como de textura para la delimitación de masas forestales a partir de imágenes de satélite. Inicialmente la imagen se clasifica basándose exclusivamente en las componentes espectrales de color, donde cada píxel se etiqueta de acuerdo a la clase a la que pertenece, utilizando la técnica de división y fusión de regiones (Pajares y Cruz, 2007; Gonzalez y Woods, 2007; Koschan y Abidi, 2008). Tras la clasificación inicial, se determinan los bordes de las regiones en base a las propiedades derivadas de los coeficientes de textura obtenidos mediante la transformada de wavelets.

En Wan y col. (2007) se combina el color y la textura utilizando los descriptores del estándar MPEG-7 para la extracción del contenido visual de las imágenes con texturas naturales.

Trias-Sanz y col. (2008) combinan de manera jerárquica las propiedades espectrales con los descriptores de texturas. En su estudio analizan diversos modelos de color combinados a partes iguales con diversos descriptores de texturas. En primer lugar obtienen una segmentación inicial mediante el color y seguidamente realizan una homogeneización de las regiones basándose en la distribución espacial de los descriptores de textura.

Ozden y Polat (2007) combinan dos tipos de propiedades: *a*) componentes espectrales de cada píxel y *b*) valores obtenidos a partir de los descriptores de textura. Todo ello considerando la información espacial proporcionada por la vecindad de cada píxel. La información procedente de los descriptores de textura la obtienen a través de los coeficientes de la transformada de wavelets (Pajares y Cruz, 2007), considerando los coeficientes de detalle obtenidos mediante dicha transformada. Para incorporar la información espacial utilizan el procedimiento conocido como *Mean-Shift* (Comaniciu y Meer, 2002).

En Chanussot y col. (2006) se propone una estrategia que utiliza, en un primer nivel, la información espectral para clasificar imágenes urbanas obtenidas mediante el satélite IKONOS (SIC, 2009). En un segundo nivel utiliza la información espacial, que en este caso es información estructural, en el sentido de que aprovecha cierta información de este tipo para mejorar la clasificación. Exactamente, establece relaciones tales como que los tejados de las casas deben tener una forma geométrica determinada, las carreteras son estructuras alargadas, etc. De este modo, un píxel incorrectamente clasificado según alguna de las clases que identifican dichas formas se puede subsanar por la consideración de este tipo de información estructural.

## **2.5 Métodos de clasificación mediante información espacial**

Aunque el tema de la incorporación de la información espacial para la clasificación de texturas ha sido tratado previamente, en esta sección se enfoca desde una perspectiva más general, abordando así las texturas espacialmente irregulares, es decir, sin patrones

predefinidos. En esto se diferencian de las texturas estudiadas en las secciones 2.4.2 y 2.4.3 donde los descriptores de textura se basan precisamente en distribuciones espaciales regulares.

Las relaciones espaciales se pueden establecer de diversas maneras. De entre todas ellas destaca el enfoque basado en los denominados *Campos Aleatorios de Markov* (*Markov Random Fields*, MRFs) (Geman y Geman, 1984; Li y col., 2000). Las relaciones espaciales pueden establecerse también en función de ciertas propiedades estructurales, por ejemplo, incorpora la información espacial mediante la inyección de conocimiento procedente de la percepción humana (Chen y col., 2005). Se parte de los niveles elementales de color, a nivel de píxel (proceso local), para posteriormente utilizar descriptores de textura que incorporan información semántica. Un ejemplo de este tipo de información se refiere a situaciones tales como que las texturas correspondientes al cielo se situarán normalmente en la parte superior de la imagen y que en las proximidades de un lago habrá texturas procedentes de zonas de vegetación.

Finalmente, otros métodos establecen relaciones espaciales sin encuadrarse bajo una perspectiva concreta, que sólo tienen en común el hecho de la incorporación de información espacial, lo cual se lleva a cabo bajo distintas perspectivas, que a su vez difieren de las expresadas anteriormente. Algunos de ellos se han utilizado en otros contextos diferentes al de las imágenes de texturas en entornos de exterior, sin embargo sirven como ejemplos adicionales para la incorporación de la información espacial (Xie y Tsui, 2004; Liew y col., 2003; Krishnapuram y Freg, 1992; Kupinski y Gigger, 1998; Precioso y col., 2005; Kim y col., 2006).

## 2.6 Resumen del análisis bibliográfico

Tras el análisis de diversos métodos, Hu y Damper (2008) concluyen que no existe ninguna estrategia combinada válida de forma general. Así, dependiendo de la aplicación, una misma estrategia de combinación puede conducir a resultados diferentes. Esta consideración, junto con la revisión de métodos realizada en este capítulo, proporciona las pautas necesarias para abordar y estudiar nuevas estrategias de combinación de clasificadores, lo que constituye el fundamento del trabajo de investigación que se presenta en esta memoria.

Tras el análisis bibliográfico realizado previamente, se llega a las siguientes conclusiones, que darán soporte a los planteamientos propuestos en este trabajo:

- Es necesario definir convenientemente las características a utilizar. Se han propuesto diferentes modelos de color, resultando la representación *RGB* una de las más apropiadas. Con la utilización de un espacio tridimensional, como es el caso del modelo *RGB*, se evitan algunos problemas relacionados con el hecho de usar espacios de alta dimensionalidad, tal y como se explica en Ho (2002). Una de las consecuencias derivadas de dicho trabajo estriba en el hecho de que a mayor dimensionalidad más difícil resulta la distribución de las muestras en clases.
- Conviene utilizar técnicas combinadas frente al uso de métodos individuales. Esto está avalado por el mejor comportamiento obtenido con los métodos combinados frente a los métodos simples, tal y como se deriva de los estudios realizados en los trabajos referenciados en este capítulo, así como de las tendencias actuales.
- Aunque los clasificadores supervisados puedan obtener mejores resultados que los no supervisados, debido a la intervención del propio supervisor, en aras de la automatización de los procesos, es preferible diseñar métodos de clasificación no supervisada por el mero hecho de tener que manejar abundantes datos.
- Las imágenes utilizadas en este trabajo de investigación no poseen patrones de textura repetitivos, ni relaciones estructurales entre objetos de la escena. Por este motivo, la combinación de clasificadores puede aplicarse de forma local, esto es píxel a píxel.
- No obstante, dadas las características espaciales de toda imagen, es posible utilizar este tipo de información global sobre la base de que dado un píxel, es probable que en su vecindad haya píxeles de su misma naturaleza, pertenecientes a la misma región de textura o similares. En caso contrario delimitará zonas de diferentes texturas. Este es uno de los enfoques ampliamente utilizado en clasificación de texturas, como se ha puesto de manifiesto en los métodos basados en el paradigma de los MRFs.

En consecuencia, la propuesta realizada en este trabajo se centra en los siguientes aspectos:

---

1. Extracción de las características espectrales de los píxeles, en este caso, sus componentes  $R$ ,  $G$  y  $B$ , generando así los vectores de características en el espacio tri-dimensional.
2. Diseño de un enfoque de clasificación no supervisado a partir de una estrategia de naturaleza supervisada.
3. Propuesta de una estrategia de combinación a nivel local, esto es píxel a píxel, utilizando este enfoque para la combinación mediante la Teoría de la Decisión Multicriterio *Fuzzy*.
4. Valoración del comportamiento de la estrategia combinada propuesta, frente a otros métodos de clasificación individuales.



# Capítulo 3

## Estrategia combinada de clasificación

### 3.1 Introducción

El presente capítulo comienza con la descripción de cinco clasificadores clásicos individuales, cuatro de naturaleza supervisada y uno de naturaleza no supervisada, que constituyen la base para el diseño de las estrategias de clasificación combinadas propuestas en este trabajo. Por clasificador *clásico* entendemos la descripción del clasificador correspondiente según su versión original. Por clasificador *individual* se entiende el hecho de utilizar un único clasificador, para distinguirlo de aquellas estrategias que emplean dos o más clasificadores, a las que nos referiremos como *combinadas*.

Los cinco clasificadores individuales, cuyos fundamentos se describen en la sección 3.2 son los siguientes (Pajares y Cruz, 2007): *agrupamiento borroso*, *clasificador paramétrico Bayesiano*, *ventana de Parzen*, *algoritmo generalizado de Lloyd* y *mapas auto-organizativos*. La elección de los mismos se basa en su probada eficacia a nivel individual en diversos ámbitos de aplicación, incluyendo la clasificación de imágenes.

Como se ha puesto de manifiesto en el capítulo dos, la combinación de clasificadores mejora los resultados, razón por la cual se elige una de las numeras opciones posibles para realizar la combinación de los cinco clasificadores individuales, optando por la *Toma de Decisiones Multicriterio bajo el paradigma de la lógica Fuzzy*, entre otras razones por su naturaleza fuzzy, que permite un cierto margen de flexibilidad respecto de las imágenes utilizadas en los experimentos llevados a cabo en este trabajo.

Aunque bien es cierto que se podían haber elegido tanto otros clasificadores individuales diferentes como una estrategia combinada distinta, la conjunción de ambos en relación a la mejora de los resultados demuestra, en el conjunto de imágenes utilizadas, sus posibilidades prometedoras. Todo ello, sin descartar en el futuro la utilización de nuevos

clasificadores, así como de otras estrategias de combinación para el tipo de imágenes analizadas.

A pesar de la naturaleza supervisada implícita en cuatro de los cinco clasificadores individuales, la propuesta combinada que se formula en el presente trabajo de investigación es de naturaleza no supervisada. Esto se consigue mediante el diseño de la estrategia propuesta, que permite distribuir las muestras en clases de forma automática, esto es, realizar una partición, junto con el correspondiente proceso de validación de dicha partición. Por tal motivo, es necesario en primer lugar, establecer el proceso para obtener la partición y en segundo lugar, fijar los criterios para su validación.

Tanto los métodos individuales como los combinados poseen una característica común, consistente en su funcionamiento en dos fases: *aprendizaje* y *decisión*.

### **3.1.1 Fases de entrenamiento y decisión**

Todo proceso de clasificación en general, y en particular la identificación de texturas en imágenes naturales, lleva asociadas dos fases fundamentales: *entrenamiento* y *decisión*. Nos referiremos a la primera fase también como de *aprendizaje*, por identificarse ambos conceptos en la literatura especializada. Por la naturaleza de su procesamiento en el tiempo a veces aparecen nombradas como procesos *off-line* y *on-line* respectivamente. Ello se debe al hecho de que la fase de entrenamiento se suele llevar a cabo durante periodos de inactividad del sistema, siendo en este momento cuando se estiman o aprenden los parámetros involucrados en el proceso. Por contra, la fase de decisión se realiza durante la plena operatividad del sistema, utilizando los parámetros aprendidos en la fase de entrenamiento.

La figura 3.1 muestra un esquema general de funcionamiento del sistema de aprendizaje/decisión particularizado para el caso de las imágenes de texturas naturales. Ambas fases constan tanto de procesos comunes como diferenciados. En efecto, los procesos de *Captura de imágenes*, *Segmentación* y *Codificación de la información* son comunes, mientras que los procesos de *Aprendizaje* y *Decisión* son diferentes. Seguidamente se describen brevemente cada uno de ellos. Posteriormente, en cada método, bien sea individual o combinado, se establece la correspondiente diferenciación.

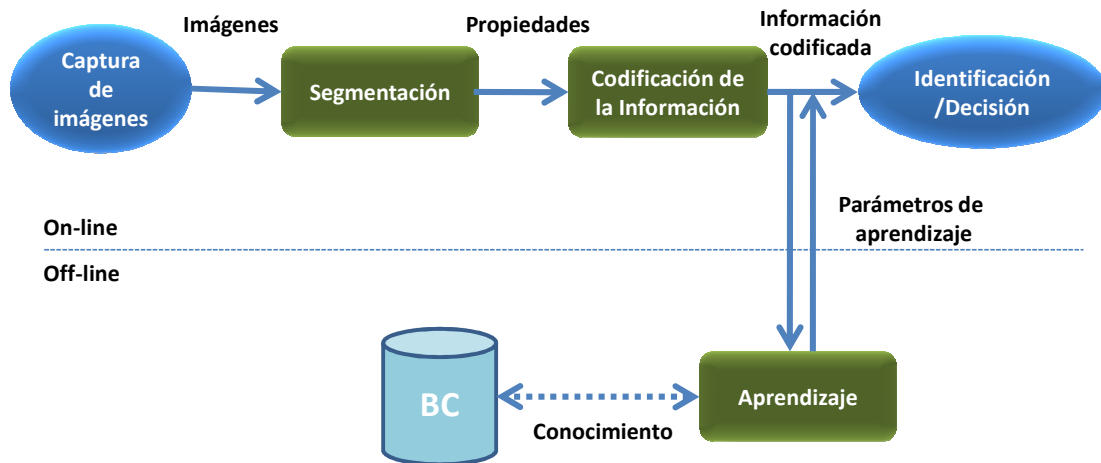


Figura 3.1 Esquema general de un proceso aprendizaje-decisión

Este esquema es válido tanto para los clasificadores de naturaleza individual como para los combinados.

- *Captura de imágenes*: consiste en la obtención de las imágenes, bien tomadas de un almacén ya existente o bien directamente a partir de la escena mediante el correspondiente sensor.
- *Segmentación*: se entiende por segmentación el procesado de tratamiento de la imagen o imágenes, consistente en la extracción de las estructuras o *características* subyacentes en ellas. Desde el punto de vista del tratamiento de imágenes, por característica se entiende algún elemento identificativo sin que necesariamente tenga que coincidir con un objeto de la escena proyectado en la imagen. Así, una característica puede ser una región o un borde que pertenece a algún objeto sin necesidad de tener que llegar a identificar dicho objeto. Una característica puede ser igualmente un píxel perteneciente a un borde, un punto de interés o sencillamente un píxel de la imagen independientemente de que pertenezca o no a alguna de las estructuras citadas. Los píxeles son las características utilizadas en este trabajo de investigación. De las características se deben extraer sus *propiedades* o *atributos* con el fin de proceder a su identificación. En el caso de una región puede ser su área, perímetro, intensidad media o cualquier otra propiedad descriptiva de la región. En nuestro caso, los atributos de los píxeles serán sus propiedades espectrales. En consecuencia, el

proceso de segmentación comprende tanto la extracción de características como sus propiedades.

- *Codificación de la información:* esta fase incluye la estructuración de la información para ser posteriormente utilizada tanto por los métodos de aprendizaje como de clasificación. Cada característica extraída durante la fase anterior, se etiqueta convenientemente, procediendo a la cuantificación de sus propiedades en un rango de variabilidad apropiado. Las etiquetas son las *muestras*, representadas por vectores, cuyas componentes son las propiedades de la característica en cuestión. Para clarificar estos conceptos y ciñéndonos al rango de nuestra aplicación, tal y como se ha mencionado previamente, las características a considerar son los píxeles. Dado un píxel en la localización espacial  $(i, j)$ , si se etiqueta como  $k$ , se tiene  $k \equiv (i, j)$ , siendo su vector representante  $\mathbf{x}_k$  cuyas componentes son los valores espectrales de dicho píxel en el modelo de color *RGB*, es decir,  $\mathbf{x}_k = \{x_{k_1}, x_{k_2}, x_{k_{31}}\} \equiv \{R, G, B\} \in \mathfrak{R}^3$  y por tanto, en este caso, el vector pertenece al espacio tridimensional  $\mathfrak{R}^3$ . Las muestras así codificadas lo son tanto para el proceso de entrenamiento como para el de decisión, distinguiéndose entre muestras de entrenamiento y muestras para la clasificación, según la fase donde se procesen.
- *Aprendizaje/Entrenamiento:* con las muestras disponibles convenientemente codificadas se procede a la realización del correspondiente proceso de entrenamiento según el método seleccionado. Los procesos de entrenamiento se describen con detalle en este capítulo. Los parámetros resultantes del aprendizaje se almacenan en la Base de Conocimiento (*BC*), figura 3.1, para su posterior uso durante la fase de decisión.
- *Identificación/Decisión:* en esta etapa se procede a identificar una nueva característica o muestra, que todavía no ha sido clasificada, como perteneciente a alguna de las clases de interés existentes. Para ello se recuperan los parámetros previamente aprendidos y almacenados en la *BC*; posteriormente y mediante la correspondiente función de decisión, inherente a cada método, se identifica la clase a la que pertenece. Este proceso también se denomina *reconocimiento* o *clasificación*. Incluso, en la literatura especializada, a veces aparece como

segmentación, si bien en el caso que nos ocupa no utilizaremos esta última denominación con el fin de diferenciarlo con nitidez del proceso de segmentación, expuesto previamente. Resulta a veces habitual el hecho de que las muestras una vez clasificadas, pueden incorporarse de nuevo al sistema, ahora ya como muestras de entrenamiento para proceder a un nuevo proceso de aprendizaje y por consiguiente llevar a cabo una actualización de los parámetros de aprendizaje asociados con cada método, que se almacenan de nuevo en la BC. Esto es lo que se conoce como aprendizaje incremental.

## 3.2 Descripción de los clasificadores individuales seleccionados

En esta sección se describen los cinco métodos de clasificación individuales mencionados en la introducción, que intervienen en el diseño del clasificador combinado propuesto en este trabajo de investigación. Cada uno de ellos se describe tanto desde el punto de vista de la fase de entrenamiento como de la fase de decisión.

### 3.2.1 Agrupamiento borroso

#### 3.2.1.1 Fase de entrenamiento

El objetivo de la técnica de agrupamiento conocida como *Agrupamiento borroso* o “*Fuzzy Clustering*” en terminología inglesa, consiste en realizar una partición de  $n$  elementos en  $c$  clases; cada uno de estos elementos es representado por el vector  $\mathbf{x}$  con  $q$  componentes, que representan las propiedades del mismo. En esta memoria se emplea el término *borroso* y *fuzzy* indistintamente.

Con el fin de formalizar el método, sea el conjunto de muestras  $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\} \in \mathfrak{R}^q$  un subconjunto del espacio real  $q$ -dimensional  $\mathfrak{R}^q$ . Cada  $\mathbf{x}_k = \{x_{k_1}, x_{k_2}, \dots, x_{k_q}\} \in \mathfrak{R}^q$  se denomina *vector de propiedades*, siendo  $x_{k_j}$  la  $j$ -ésima propiedad o atributo de la muestra  $\mathbf{x}_k$ .

Como se ha mencionado previamente, y centrándonos en las imágenes que se utilizan en este trabajo de investigación, las características son los píxeles y sus propiedades las componentes espectrales en el modelo de color *RGB*, por tanto  $q = 3$  y

$\mathbf{x}_k = \{R, G, B\} \in \mathfrak{R}^3$ . Desde el punto de vista del clasificador, se habla de clases y muestras pertenecientes a las clases, en nuestro contexto de las imágenes naturales pertenecientes a entornos de exterior, las muestras son las características, es decir, los píxeles, mientras que las clases se identifican con las texturas que aparecen en las imágenes, cuyos detalles se explican en el capítulo cuatro.

Este clasificador ha sido ampliamente utilizado en la literatura, cuya descripción se puede encontrar en Bezdek (1981), Zimmerman (1991), Duda y col., (2001) o Pajares y Cruz (2007), entre otros. A continuación se proporciona una síntesis del mismo.

El objetivo de cualquier proceso de entrenamiento debe orientarse a conseguir que las muestras se distribuyan en clases, de tal forma que las pertenecientes a una misma clase sean tan similares entre sí como sea posible y a la vez diferenciarse de las muestras de otras clases, también tanto como sea posible. Por este motivo es necesario definir alguna medida de similitud, que en el caso que nos ocupa serán funciones de distancia, de forma que la distancia entre dos muestras  $\mathbf{x}_k$  y  $\mathbf{x}_l$  determine el grado de similitud entre ellas.

La distribución de las muestras del conjunto  $X$  en clases se conoce como *partición*. Si se pretende realizar una partición del conjunto  $X$  en  $c$  clases, se tendrán  $w_i \{i = 1, \dots, c\}$  subconjuntos de  $X$ , donde cada subconjunto constituirá una clase. Una partición puede contemplarse desde las dos perspectivas siguientes: *fuzzy* y *no fuzzy*. Una partición *no fuzzy* se conoce en terminología inglesa como “*crisp*”. Desde el punto de vista *fuzzy*, para cada clase  $w_i$  se puede definir un conjunto borroso  $\mu_i$  sobre el universo  $X$ ,  $\mu_i : X \rightarrow [0,1]$  que asigna a cada muestra  $\mathbf{x}_k$  lo que se conoce como *grado de pertenencia* de dicha muestra al subconjunto  $w_i$ , denominado *clase* a partir de este momento. El grado de pertenencia de  $\mathbf{x}_k$  a la clase  $w_i$  se expresa como  $\mu_i(\mathbf{x}_k) \equiv \mu_{ik}$ . En el caso de conjuntos no borrosos o *crisp*, una muestra  $\mathbf{x}_k$  se dice que pertenece a la clase  $w_i$  dada y no pertenece al resto de clases. Esto se expresa mediante los valores discretos  $\{0,1\}$  de la siguiente forma  $\mu_{ik} = 1$  para indicar que pertenece y  $\mu_{ik} = 0$  para expresar que no pertenece. Por el contrario, en el caso de conjuntos borrosos se dice que una muestra puede pertenecer a diferentes clases y así se habla por ejemplo de que  $\mathbf{x}_k$  pertenece a una clase  $w_i$  con grado de pertenencia  $\mu_{ik}$  y a la vez a la clase  $w_j$  con grado de pertenencia

$\mu_{jk}$ . Como ejemplo, supongamos que se tienen tres clases  $w_i$ ,  $w_j$  y  $w_h$ , en este caso podríamos decir que la muestra  $\mathbf{x}_k$  pertenece a dichas clases con los siguientes grados de pertenencia  $\mu_{ik} = 0.4$ ,  $\mu_{jk} = 0.5$  y  $\mu_{hk} = 0.1$ , donde estos valores pertenecen al intervalo continuo  $[0,1]$ .

Dado el conjunto de muestras  $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  y el conjunto  $V_{cn}$  de todas las matrices reales de dimensión  $c \times n$ , con  $2 \leq c < n$ , se puede obtener una matriz de grados de pertenencia representando una partición de las muestras en clases de la siguiente manera  $U = \{\mu_{ik}\} \in V_{cn}$ . Tanto en el supuesto *crisp* como en el *fuzzy* se deben cumplir las siguientes condiciones:

$$\begin{aligned} a) \quad & \mu_{ik} \in \{0,1\} \text{ crisp o } \mu_{ik} \in [0,1] \text{ fuzzy} \quad 1 \leq i \leq c; 1 \leq k \leq n \\ b) \quad & \sum_{i=1}^c \mu_{ik} = 1 \quad 1 \leq k \leq n \\ c) \quad & 0 < \sum_{k=1}^n \mu_{ik} < n \quad 1 \leq i \leq c \end{aligned} \tag{3.1}$$

Para ilustrar los conceptos anteriores, sea  $X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ , entonces podemos construir las siguientes matrices suponiendo que  $c = 2$ , es decir, se dispone de dos clases. En ambas matrices las filas se asocian con las clases, así si nos fijamos en  $U_1$  diríamos que  $\mathbf{x}_1$  pertenece a la clase  $w_1$  con grado de pertenencia uno y no pertenece a la clase  $w_2$  al tener un grado de pertenencia cero. Por el contrario, si consideramos  $U_2$ , ahora  $\mathbf{x}_1$  pertenece a  $w_1$  con grado de pertenencia 0.9 pero simultáneamente, también pertenece a  $w_2$  con grado de pertenencia 0.1.

$$\begin{array}{cc} \text{crisp} & \text{fuzzy} \\ U_1 = \begin{matrix} w_1 & \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \\ w_2 & \end{matrix} & U_2 = \begin{matrix} w_1 & \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \\ 0.9 & 0.4 & 0.2 \\ 0.1 & 0.6 & 0.8 \end{bmatrix} \\ w_2 & \end{matrix} \end{array}$$

Una clase  $w_j$  se representa por su centro  $\mathbf{v}_j = \{v_{j_1}, v_{j_2}, \dots, v_{j_q}\} \in \Re^q$  teniendo en cuenta que  $j = 1, \dots, c$ . Intuitivamente, este centro sería el representante de las muestras en la clase,

las cuales tienden a agruparse a su alrededor. Como se ha mencionado previamente en el caso de las imágenes utilizadas en este trabajo,  $q$  es igual a tres.

El proceso de entrenamiento o aprendizaje recibe las muestras del conjunto  $X$ , siendo conocido el número de clases,  $c$ . Las muestras se encuentran distribuidas en esas  $c$  clases. Dado que es necesario el conocimiento de  $c$ , el proceso se engloba dentro de los métodos de clasificación de naturaleza *supervisada*. El objetivo final de este proceso consiste en estimar los centros de las clases así como los grados de pertenencia de cada muestra a cada clase. Con tal propósito se define la función objetivo  $J$ , que debe ser minimizada asumiendo las restricciones definidas en la ecuación (3.1) (Bezdek, 1981; Duda y col., 2001; Zimmerman, 1991),

$$J(U; \mathbf{v}) = \sum_{i=1}^n \sum_{j=1}^c \mu_{ij}^m d_{ij}^2 \quad (3.2)$$

donde  $\mathbf{v} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_c\}$ ,  $\mathbf{v}_j \in \mathbb{R}^q$  representa el centro de la clase  $j$ -ésima;  $d_{ij}^2 \equiv d^2(\mathbf{x}_i, \mathbf{v}_j)$  expresa una medida de distancia al cuadrado entre la muestra  $\mathbf{x}_i$  y el centro de la clase  $w_j$  dado por  $\mathbf{v}_j$ ;  $m$  se conoce como *peso exponencial* (Bezdek, 1981).

Para minimizar la función objetivo, los centros de las clases y los grados de pertenencia conviene elegirlos de modo que los valores altos en los grados de pertenencia se obtengan para aquellas muestras que estén cerca, en términos de distancia, de los correspondientes centros de las clases. Por otro lado, cuanto mayor sea el valor de  $m$ , menor será la contribución de las muestras con valores bajos de los grados de pertenencia a la función objetivo. Dichas muestras tienden a ser ignoradas en la estimación de los centros de las clases y grados de pertenencia (Zimmerman, 1991). El parámetro  $m$  se ajusta durante la fase experimental mediante el procedimiento de validación cruzada (“*cross-validation*”) descrito en el capítulo cuatro, sección 4.2.1 (Duda y col., 2001).

El algoritmo del clasificador borroso se fundamenta en un proceso iterativo, de forma que en cada iteración  $t$ , se modifican los centros de las clases en función de los grados de pertenencia que también se actualizan, según las ecuaciones (3.3) y (3.4),

$$\mu_{ij}(t+1) = \frac{1}{\sum_{r=1}^c (d_{ij}(t)/d_{ir}(t))^{2/(m-1)}} \quad (3.3)$$



$$\mathbf{v}_j(t+1) = \frac{\sum_{i=1}^n \mu_{ij}^m(t) \mathbf{x}_i}{\sum_{i=1}^n \mu_{ij}^m(t)} \quad (3.4)$$

El proceso se detiene en aquella iteración en la que se verifica que  $\|\mu_{ij}(t+1) - \mu_{ij}(t)\| < \varepsilon \quad \forall ij$ , lo que se conoce como *criterio de convergencia*. No obstante, en la práctica se suele establecer un criterio adicional fijando un número máximo de iteraciones  $t_{max}$ , para que el proceso se detenga en caso de no conseguir la convergencia.

Los centros de las clases constituyen los parámetros de aprendizaje, que se almacenan en la *BC*, figura 3.1.

El proceso iterativo descrito previamente requiere, o bien comenzar con unos grados de pertenencia dados y a partir de ellos calcular los centros o a la inversa, esto es, comenzar con unos centros dados para posteriormente continuar calculando los grados de pertenencia. En este trabajo se ha optado por la segunda alternativa. Con tal propósito, se ha elegido el método pseudo-aleatorio propuesto por Balasko y col. (2006), gracias al cual se inicializan los centros calculando las medias estadísticas de las clases, obviamente suponiendo conocido el número de clases. De este modo, el proceso de convergencia comienza con unos valores de los centros próximos a los finales. La inicialización de los centros se lleva a cabo según los dos pasos siguientes:

1. Realizar una transformación lineal  $Y = f(X)$  de las muestras de entrenamiento  $X$ , de tal forma que sus valores queden restringidos al rango  $[0,1]$ .
2. Inicializar el vector de centros  $\mathbf{v} = 2D\bar{\mathbf{M}} \circ \mathbf{R} + D\bar{\mathbf{m}}$ , donde  $\bar{\mathbf{m}}$  es el vector media para la transformada de las muestras de entrenamiento con los valores de  $Y$  y  $\bar{\mathbf{M}} = \max(\text{abs}(Y - \bar{\mathbf{m}}))$ , ambos del tamaño  $1 \times d$ ;  $D = [1 \quad \dots \quad 1]^T$  con tamaño  $c \times 1$ ;  $\mathbf{R}$  es una matriz de número aleatorios con valores en el intervalo  $[0,1]$  y tamaño  $c \times q$ ; la operación  $\circ$  define la multiplicación elemento por elemento de las matrices.

### 3.2.1.2 Fase de decisión

La decisión constituye el proceso por el cual una nueva muestra  $\mathbf{x}_s$  cuya pertenencia a las clases resulta desconocida hasta el momento, debe ser identificada como perteneciente a una de las clases  $w_j$  disponibles. Con tal propósito, se recuperan de la  $BC$ , figura 3.1, los centros de las clases  $\mathbf{v}_j$  ( $j = 1, 2, \dots, c$ ) a partir de los cuales se obtienen los grados de pertenencia de  $\mathbf{x}_s$  a cada una de las clases según la siguiente expresión,

$$\mu_{sj} = \frac{1}{\sum_{r=1}^c (d_{sj} / d_{sr})^{2/(m-1)}} \quad (3.5)$$

donde  $d_{sj}$  define, como antes, una distancia entre la muestra  $\mathbf{x}_s$  y el centro de la clase  $\mathbf{v}_j$ ,  $d_{sj} \equiv d(\mathbf{x}_s, \mathbf{v}_j)$ ; en el caso que nos ocupa se ha elegido la distancia Euclídea. Los grados de pertenencia dados por la ecuación (3.5) constituyen las salidas de este clasificador.

De esta forma se obtiene la siguiente secuencia de grados de pertenencia para  $\mathbf{x}_s$ ,

$$\mu_{s1} \equiv \mu(\mathbf{x}_s, w_1), \mu_{s2} \equiv \mu(\mathbf{x}_s, w_2), \dots, \mu_{sc} \equiv \mu(\mathbf{x}_s, w_c) \quad (3.6)$$

Finalmente,  $\mathbf{x}_s$  se clasifica como perteneciente a la clase  $w_j$  que proporcione el máximo grado de pertenencia según la siguiente regla de decisión,

$$\mathbf{x}_s \in w_j \mid \mu_{sj} > \mu_{sk} \quad \forall k \neq j \text{ con } k, j = 1, 2, \dots, c \quad (3.7)$$

## 3.2.2 Clasificador Paramétrico: Bayesiano

### 3.2.2.1 Fase de entrenamiento

Este método tradicionalmente se ha venido encuadrando dentro de las técnicas de clasificación *no supervisada* (Escudero, 1977). Dada una muestra de entrenamiento genérica  $\mathbf{x} \in \mathfrak{R}^q$ , el objetivo consiste en calcular las probabilidades de pertenencia a cada una de las clases  $w_j$ , esto es  $P(w_j \mid \mathbf{x})$ .

Esta técnica parte de la base de que se conoce la función de densidad de probabilidad condicional para cada clase, resultando desconocidos los parámetros o estadísticos

involucrados en dicha función. Una práctica muy generalizada, adoptada en este trabajo, consiste en asumir que la forma de estas funciones sigue la ley de distribución Gaussiana o Normal, según la siguiente expresión,

$$p(\mathbf{x} | \mathbf{m}_j, C_j) = \frac{1}{(2\pi)^{q/2} |C_j|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mathbf{m}_j)^T C_j^{-1} (\mathbf{x} - \mathbf{m}_j) \right\} \quad (3.8)$$

donde  $\mathbf{m}_j$  y  $C_j$  son respectivamente la media y la matriz de covarianza de la clase  $w_j$ , o dicho de otra forma, los estadísticos o parámetros desconocidos que deben ser estimados;  $T$  denota la matriz transpuesta y  $q$  expresa la dimensionalidad de los datos según  $\mathbf{x} \in \mathbb{R}^q$ .

Las hipótesis que, específicamente, asumen las técnicas de clasificación no supervisadas son las siguientes:

1. Existen  $c$  clases en el problema.
2. La muestra  $\mathbf{x}$  proviene de estas  $c$  clases, aunque se desconoce la clase específica a la que pertenece.
3. La probabilidad a priori de que la muestra pertenezca a la clase  $w_j$ ,  $P(w_j)$  resulta en principio desconocida.
4. La función de densidad asociada a cada clase posee una forma conocida, siendo desconocidos los parámetros de dicha función.

Con este planteamiento resulta factible aplicar la regla de Bayes para obtener la probabilidad condicional de que  $\mathbf{x}$  pertenezca a la clase  $w_j$ , mediante la siguiente expresión (Huang y col., 2002),

$$P(w_j | \mathbf{x}) = \frac{p(\mathbf{x} | \mathbf{m}_j, C_j) P(w_j)}{\sum_{j=1}^c p(\mathbf{x} | \mathbf{m}_j, C_j)} \quad (3.9)$$

Conocidas las formas de las funciones de densidad de probabilidad, el método *paramétrico Bayesiano* o de *máxima verosimilitud* tiene por objetivo estimar los mejores parámetros para dichas funciones.

En términos de probabilidad, si  $p(\mathbf{x} | \mathbf{m}_j, C_j)$  es la probabilidad de que la muestra representada por su vector  $\mathbf{x}$  tome en la clase  $w_j$  dicho valor, entonces, teniendo en cuenta los estadísticos de todas las clases, la probabilidad de la muestra  $\mathbf{x}$  será la suma de todas las probabilidades para las  $c$  clases, que viene dada por la siguiente expresión,

$$P(\mathbf{x} | \mathbf{m}_j, C_j) = \sum_{j=1}^c p(\mathbf{x} | \mathbf{m}_j, C_j) P(w_j) \quad (3.10)$$

donde  $P(\mathbf{x} | \mathbf{m}_j, C_j)$  se conoce como probabilidad total (“*mixture density*”), las probabilidades condicionales  $p(\mathbf{x} | \mathbf{m}_j, C_j)$  son los componentes de la probabilidad total (“*component densities*”) y las probabilidades a priori son los parámetros de la mezcla (“*mixing parameters*”).

Considerando la probabilidad total, la probabilidad de que se extraiga la muestra  $\mathbf{x}$  del colectivo viene dada por la probabilidad conjunta de cada muestra  $\mathbf{x}_i$ , según la expresión,

$$P(\mathbf{x} | \mathbf{m}_j, C_j) = \prod_{i=1}^n p(\mathbf{x}_i | \mathbf{m}_j, C_j) \quad (3.11)$$

siendo  $n$  el número total de muestras disponibles en el conjunto  $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\} \in \Re^q$ .

La estimación más acertada para los parámetros  $\mathbf{m}_j$  y  $C_j$ , según el método de máxima verosimilitud, será aquella que maximice la función dada en (3.11). Desde un punto de vista analítico es preferible trabajar con el logaritmo neperiano de dicha función, ya que al ser ésta una función monótona creciente, el vector que maximiza el logaritmo también maximiza la función de (3.11). Por tanto, las mejores estimaciones para  $\mathbf{m}_j$  y  $C_j$  se obtienen calculando las primeras derivadas del logaritmo de la función dada en (3.11) e igualándolas a cero, como condición de máximo.

En la expresión (3.11) se supone que las muestras en una clase  $w_h$  no proporcionan información sobre la estimación de los parámetros para la clase  $w_j$ , lo cual significa que los parámetros para las diferentes clases son funcionalmente independientes. Esto se traduce en que es posible trabajar con cada clase separadamente y en consecuencia el número de muestras total  $n$  en (3.11) puede sustituirse por  $n_j$ , que representa el número de muestras pertenecientes a la clase  $w_j$ , que intervienen en el cómputo de la función de

densidad de probabilidad asociada a dicha clase. Por tanto, de aquí en adelante, y bajo la suposición de que se tienen  $c$  problemas, el número de muestras que intervienen en el cómputo de los parámetros de la clase  $w_j$  será  $n_j$  (Duda y col., 2001).

Consideremos el supuesto univariable con  $\mathbf{m}_j = m_j$  y  $C_j = \sigma_j^2$  y  $\mathbf{x}_i = x_i$ , en cuyo caso,

$$\ln P(\mathbf{x} | \mathbf{m}_j, C_j) = \sum_{i=1}^{n_j} \ln p(x_i | m_j, C_j) \quad (3.12)$$

teniendo en cuenta la ecuación (3.8) para el caso unidimensional,  $q = 1$ , resulta,

$$\ln p(x_i | m_j, C_j) = -\frac{1}{2} \ln 2\pi C_j - \frac{1}{2C_j} (x_i - m_j)^2 \quad (3.13)$$

Derivando la expresión (3.13) con respecto a cada uno de los parámetros que se están estimando, es decir,  $m_j$  y  $C_j$  se tiene,

$$\frac{\partial}{\partial m_j} \ln p(x_i | m_j, C_j) = \frac{1}{C_j} (x_i - m_j) \quad (3.14)$$

$$\frac{\partial}{\partial C_j} \ln p(x_i | m_j, C_j) = -\frac{1}{2C_j} + \frac{(x_i - m_j)^2}{2C_j^2} \quad (3.15)$$

Considerando las  $n_j$  muestras de entrenamiento, de acuerdo con la discusión previa y según la ecuación (3.12), se obtiene,

$$\frac{1}{n_j} \sum_{i=1}^{n_j} \frac{1}{\hat{C}_j} (x_i - \hat{m}_j) = 0 \quad -\frac{1}{n_j} \sum_{i=1}^{n_j} \frac{1}{2\hat{C}_j} + \frac{1}{n_j} \sum_{i=1}^{n_j} \frac{(x_i - \hat{m}_j)^2}{2\hat{C}_j^2} = 0 \quad (3.16)$$

donde  $\hat{m}_j$  y  $\hat{C}_j$  son las estimas de máxima verosimilitud para  $m_j$  y  $C_j$ , respectivamente.

Resolviendo, obtenemos finalmente las estimaciones para los parámetros desconocidos de acuerdo con la siguiente expresión,

$$\hat{m}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_i \quad \hat{\sigma}_j^2 = \frac{1}{n_j} \sum_{i=1}^{n_j} (x_i - \hat{m}_j)^2 \quad (3.17)$$

Aunque el análisis del caso multivariable es básicamente similar, se requiere mucha más manipulación para su obtención. El resultado, muy bien conocido en estadística, es que las estimaciones de máxima verosimilitud para  $\mathbf{m}_j$  y  $C_j$  vienen dadas por,

$$\mathbf{m}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} \mathbf{x}_i \quad C_j = \frac{1}{n_j - 1} \sum_{i=1}^{n_j} (\mathbf{x}_i - \mathbf{m}_j)(\mathbf{x}_i - \mathbf{m}_j)^T \quad (3.18)$$

La ecuación (3.18) expresa que la estima de máxima verosimilitud para el vector media es la media aritmética simple, mientras que la estima de máxima verosimilitud para la matriz de covarianza es la media aritmética de las  $n_j$  matrices  $(\mathbf{x}_i - \mathbf{m}_j)(\mathbf{x}_i - \mathbf{m}_j)^T$ .

A partir de este momento quedan totalmente especificadas las funciones de densidad de probabilidad definidas en la ecuación (3.8), ya que los parámetros, que allí eran desconocidos, acaban de ser estimados. Estos parámetros  $\mathbf{m}_j$  y  $C_j$  ( $j = 1, 2, \dots, c$ ) se almacenan finalmente en la *BC* de acuerdo con el esquema de la figura 3.1.

### 3.2.2.2 Fase de decisión

Dada una nueva muestra  $\mathbf{x}_s$ , el objetivo consiste en clasificarla como perteneciente a una de las clases disponibles. Para ello se calculan las probabilidades condicionales de que  $\mathbf{x}_s$  pertenezca a cada una de las clases  $w_j$  de acuerdo con la expresión (3.9) según la siguiente regla de decisión (Duda y col., 2001),

$$\mathbf{x}_s \in w_j \quad \text{si} \quad P(w_j | \mathbf{x}_s) > P(w_k | \mathbf{x}_s) \quad \forall k \neq j \quad (3.19)$$

No obstante, en esta regla el denominador de la ecuación (3.9) no intervine en la decisión por ser idéntico en ambos miembros de la desigualdad; por consiguiente, la regla de decisión puede reescribirse como sigue,

$$\mathbf{x}_s \in w_j \quad \text{si} \quad p(\mathbf{x}_s | \mathbf{m}_j, C_j) P(w_j) > p(\mathbf{x}_s | \mathbf{m}_k, C_k) P(w_k) \quad \forall k \neq j \quad (3.20)$$

La única cuestión pendiente estriba en el cómputo de las probabilidades a priori. Cuando éstas no se conocen, normalmente la forma de proceder consiste en fijar todas las probabilidades a priori al mismo valor, por ejemplo a  $1/2$ , en cuyo caso, dejarían de

participar en la decisión final. Esta es la aceptación hecha en este trabajo, por lo que la expresión (3.20) se transforma en la (3.21),

$$\mathbf{x}_s \in w_j \quad \text{si} \quad p_b(\mathbf{x}_s | w_j) > p_b(\mathbf{x}_s | w_k) \quad \forall k \neq j \quad (3.21)$$

No obstante, en algunas estrategias combinadas donde participa este clasificador, es necesario normalizar los valores de probabilidad al rango  $[0, +1]$  para competir en igualdad de condiciones con el resto de clasificadores que intervienen en la combinación. Por este motivo, en sus actuaciones combinadas, la regla de decisión es la proporcionada por la ecuación (3.19), ya que gracias al efecto del denominador del segundo miembro de la ecuación (3.9) el resultado final está normalizado a dicho rango.

Por consiguiente y según lo anterior, en este trabajo de investigación, las salidas del clasificador son las probabilidades proporcionadas a través de la ecuación (3.9), si bien renombradas con el subíndice  $b$ , esto es  $P_b(w_j | \mathbf{x})$  para diferenciarlas posteriormente de las probabilidades que se obtienen a través del estimador no paramétrico basado en la ventana de Parzen.

### 3.2.3 Estimación no paramétrica: ventana de Parzen

#### 3.2.3.1 Fase de entrenamiento

En este proceso, al igual que en el caso paramétrico de máxima verosimilitud, el objetivo sigue siendo el cálculo de las probabilidades de pertenencia de la muestra  $\mathbf{x}$  a cada una de las clases  $w_j$ , esto es  $P(w_j | \mathbf{x})$ .

Por tanto, el problema se plantea desde el mismo punto de vista, formulando las mismas tres primeras hipótesis y sustituyendo la cuarta por una nueva que ahora se concreta como: “la forma de la función de densidad de probabilidad asociada a cada clase no se conoce”. Esto significa que en este caso no existen parámetros a estimar, siendo por tanto la propia función de densidad de probabilidad el objetivo de la estimación (Parzen, 1962, Duda y col. 2001). La estimación se lleva a cabo considerando la influencia de cada una de las muestras en la función de densidad de probabilidad a través de una función potencial, que determina el grado de influencia en la participación de cada muestra, tal y como se expresa en la ecuación (3.22) con  $n_j$  el número de muestras en la clase  $w_j$ ,

$$p_p(\mathbf{x} | w_j) = \frac{1}{n_j} \sum_{k=1}^{n_j} K_h(\mathbf{x} - \mathbf{x}_k) \quad (3.22)$$

En la ecuación anterior,  $K_h(\mathbf{z}) = K(\mathbf{z}/h)/V_h$ , define la función potencial a través de la denominada ventana o función del núcleo  $K(\cdot)$ , cumpliendo  $\int_{-\infty}^{+\infty} K(k)dk = 1$ . Esta propiedad establece que la función de densidad de probabilidad cumple que el área bajo ella debe ser igual a la unidad;  $h$  se conoce como parámetro de suavizado o ancho de la ventana, que determinará el intervalo de influencia del núcleo, con  $h > 0$  (Fukunaga y Hummels, 1987). El subíndice  $p$  hace referencia a la probabilidad de Parzen, para distinguirla de la probabilidad basada en la estimación de máxima verosimilitud definida previamente e identificada con el subíndice  $b$ .

Típicamente, la función potencial se elige de forma que sea una función monótona decreciente, tal que disminuya con la distancia de su argumento a la muestra  $\mathbf{x}_j$ . Si llamamos a esta función decreciente  $g$ , entonces  $K(\cdot)$  se puede expresar a través de  $g$  de forma que:  $K(\mathbf{x} - \mathbf{x}_k) = g(d(\mathbf{x}, \mathbf{x}_k))$ , donde  $d(\mathbf{x}, \mathbf{x}_j)$  es una distancia métrica. Las propiedades estadísticas para este estimador se pueden encontrar en Duda y col. (2001).

La ecuación (3.22) expresa que la contribución de cada muestra de entrenamiento  $\mathbf{x}_k$ , a la estimación de la función de densidad en el punto  $\mathbf{x}$  está determinada por el ancho y la forma de la función del núcleo. Por ejemplo, si la ventana tiene la forma Gaussiana, las muestras que caen cerca de  $\mathbf{x}$  contribuyen más a la estimación de la función de densidad que aquellas otras que se sitúan lejos de  $\mathbf{x}$ . Tales ecuaciones son interpretadas como un proceso de centrado del núcleo  $K(\cdot)$  sobre cada muestra de entrenamiento, posteriormente se suman las  $n_j$  funciones solapadas en el espacio alrededor de  $\mathbf{x}$ , de ahí el sumatorio que aparece en la ecuación (3.22).

A la hora de estimar la función de densidad de probabilidad, la forma del núcleo no es tan importante como su ancho. Se han propuesto diferentes tipos de núcleos (Duda y col., 2001). En este trabajo se ha elegido el núcleo Gaussiano, por las dos razones que se exponen a continuación. La primera es porque con este núcleo obtenemos toda la información sobre el grado de correlación entre los valores de las propiedades de las muestras de entrenamiento a través de la matriz de covarianza  $C_j$ , para la clase  $w_j$ . La



segunda razón viene dada por sus excelentes propiedades analíticas. En consecuencia, la función de densidad finalmente estimada resulta ser la proporcionada por la ecuación (3.23).

$$p_p(\mathbf{x} | w_j) = \frac{1}{n_j} \sum_{k=1}^{n_j} \left\{ \frac{\exp\{-D(\mathbf{x}, \mathbf{x}_k, h_j)\}}{(2\pi)^{q/2} h_j^{n_j} |C_j|^{1/2}} \right\} \quad (3.23)$$

donde  $D(\cdot) = (\mathbf{x} - \mathbf{x}_k)^T C_j^{-1} (\mathbf{x} - \mathbf{x}_k) / 2h_j^2$ ;  $q$  representa la dimensión de las muestras en el espacio considerado;  $T$  indica la operación de transposición de vectores. Cada función de núcleo considera que las muestras se sitúan alrededor del patrón teórico  $\mathbf{x}_k$  donde el vector promedio para cada núcleo es el propio patrón. El parámetro de suavizado  $h_j$  se expresa generalmente en función del número de patrones en la clase, esto es  $n_j$ ,

$$h_j = h n_j^{-\frac{r}{q}} \quad \text{para} \quad 0 \leq r \leq 0.5 \quad (3.24)$$

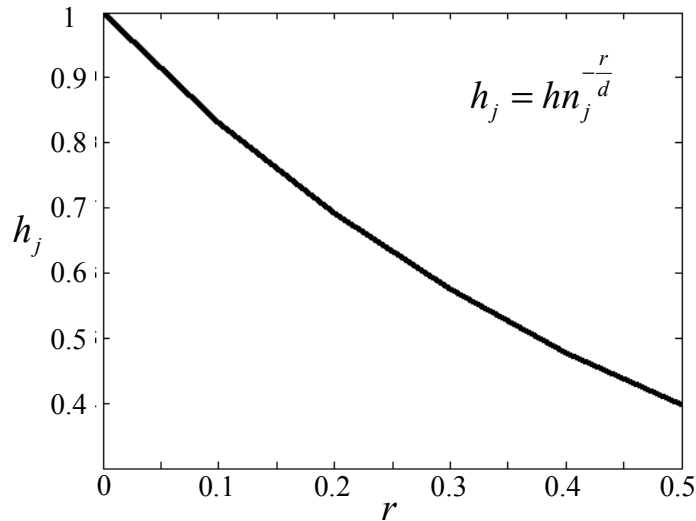
La elección del ancho de banda  $h$  es muy crítica en la estimación de la función de densidad mediante Parzen (Duin, 1976). En efecto, valores pequeños de  $h$  producen estimas con picos significativos centrados sobre las muestras, correspondiéndose con la posición de los núcleos sobre dichas muestras. Por el contrario, cuando  $h$  es muy grande cada patrón realiza la misma contribución en cada punto  $\mathbf{x}$ ; el resultado, en este caso, es una sobre-estima de  $p_p(\mathbf{x} | w_j)$ . Como se indica en Duin (1976) el ancho de la ventana debe elegirse de acuerdo con el número de muestras procesadas. Una discusión más en profundidad sobre la elección del ancho de la ventana se puede encontrar en los trabajos de Jain y Ramaswami (1988) o Parzen (1962). En efecto, en el primero aparece un detallado estudio para el caso bi-dimensional con un núcleo Gaussiano para valores de  $n_j$  y  $h$ . En dicho estudio se encuentran como valores aceptables los obtenidos con  $n_j = 256$  y  $h = 1$ , si bien los mejores resultados se obtienen para  $n_j \rightarrow +\infty$  sin importar el valor de  $h$ . A partir de aquí y con el fin de determinar el mejor valor de  $h$ , en este trabajo se ha diseñado una estrategia, que se concreta en los siguientes pasos,

1. Fijar  $n_j$  a 256 y  $h$  igual a 1.
2. Variar  $r$  desde 0.0 hasta 0.5 con incrementos de 0.05.

3. Obtener los valores de  $h_j$  en función de los valores de  $r$  según la ecuación (3.24)
4. Como durante la fase de entrenamiento el número de muestras disponibles resulta conocido, siendo exactamente  $n_j$ , se vuelve a variar  $r$  como antes desde 0.0 hasta 0.5 con incrementos de 0.05; de este modo, por medio de la ecuación (3.24) se obtiene una secuencia de valores de  $h = h_j n_j^{r/q}$  donde para cada  $r$ ,  $h_j$  es el obtenido en el paso tres.
5. La única cuestión pendiente consiste en determinar el mejor valor de  $r$  y por tanto de  $h$ , ya que  $n_j$  viene fijado por el número de muestras de la clase y  $h_j$  se calcula mediante los valores obtenidos en el paso tres; el mejor valor de  $r$  buscado se determina a partir de los datos disponibles mediante validación cruzada (Duda y col., 2001), tal y como se explica en 4.2.1.

En la figura 3.2 se representa la variación de  $h_j$  frente a  $r$  par los valores de  $h = 1$  y  $n_j = 256$ .

De acuerdo con la ecuación (3.23), este clasificador estima la función de densidad de probabilidad teniendo en cuenta las muestras de entrenamiento asociadas a cada clase, necesitando que las muestras estén ya distribuidas en ellas, o lo que es equivalente, la partición debe estar disponible; además utiliza las matrices de covarianza asociadas a cada una de las clases. La partición completa y las matrices de covarianza son los parámetros que este clasificador almacena en la  $BC$ , figura 3.1. En realidad, las matrices de covarianza son las mismas que las obtenidas por el clasificador de máxima verosimilitud, no siendo necesario su cálculo nuevamente.


 Figura 3.2 Variación de  $h_j$  frente a  $r$  para valores dados de  $h$  y  $n_j$ 

### 3.2.3.2 Fase de decisión

Durante la fase de decisión el clasificador de Parzen extrae de la  $BC$  tanto las matrices de covarianza  $C_j$  como las muestras de entrenamiento disponibles y distribuidas en sus correspondientes clases. Con ellas se genera la función de densidad de probabilidad dada en (3.23). De este modo, ante la llegada de una nueva muestra  $\mathbf{x}_s$  se obtienen las probabilidades condicionales según dicha ecuación,  $p_p(\mathbf{x}_s | w_j)$ . La probabilidad de que la muestra  $\mathbf{x}_s$  pertenezca a la clase  $w_j$  puede obtenerse aplicando de nuevo la regla de Bayes,

$$P_p(w_j | \mathbf{x}_s) = \frac{p_p(\mathbf{x}_s | w_j) P(w_j)}{\sum_{j=1}^c p_p(\mathbf{x}_s | w_j)} \quad (3.25)$$

Al igual que en el caso de la estimación por máxima verosimilitud, es necesario conocer la probabilidad a priori. No obstante, en el supuesto de que no se conozca se asume que toma el mismo valor para todas las clases, por ejemplo 0.5; esta es la suposición hecha en este trabajo de investigación, al igual que en el caso del estimador de máxima verosimilitud, como ya se indicó previamente.

La clasificación de  $\mathbf{x}_s$  se determina según la siguiente regla de decisión (Duda y col., 2001),

$$\mathbf{x}_s \in w_j \quad \text{si} \quad P(w_j | \mathbf{x}_s) > P(w_k | \mathbf{x}_s) \quad \forall k \neq j \quad (3.26)$$

Debido a que el denominador en la ecuación (3.25) es el mismo para todas las clases, éste no interviene en la decisión, por lo que la regla (3.26) puede reescribirse como sigue,

$$\mathbf{x}_s \in w_j \quad \text{si} \quad p(\mathbf{x}_s | w_j)P(w_j) > p(\mathbf{x}_s | w_k)P(w_k) \quad \forall k \neq j \quad (3.27)$$

Como en el estimador de máxima verosimilitud, cuando las probabilidades a priori no se conocen de antemano, siendo este nuestro caso, la regla de decisión resulta ser finalmente,

$$\mathbf{x}_s \in w_j \quad \text{si} \quad p_p(\mathbf{x}_s | w_j) > p_p(\mathbf{x}_s | w_k) \quad \forall k \neq j \quad (3.28)$$

Al igual que en el caso del estimador de máxima verosimilitud y debido a la necesidad de restringir los valores de probabilidad al rango  $[0,1]$  cuando el clasificador actúa combinado con otros clasificadores, la regla de decisión utilizada es la definida en (3.26).

Las salidas del clasificador son las probabilidades dadas por la ecuación (3.25), esto es  $P_p(w_j | \mathbf{x}_s)$ , donde el subíndice  $p$  las distingue de las probabilidades de salida proporcionadas por el clasificador basado en máxima verosimilitud, que se distinguen con el subíndice  $b$ , como se ha mencionado previamente.

### 3.2.4 Algoritmo Generalizado de Lloyd

#### 3.2.4.1 Fase de entrenamiento

A continuación vamos a describir el algoritmo generalizado de Lloyd (AGL) propuesto inicialmente por Lloyd (1957) y generalizado más tarde para cuantización vectorial por Linde y col. (1980). Este algoritmo aplica las dos condiciones necesarias a los datos de entrenamiento para determinar cuantizadores vectoriales óptimos minimizando el riesgo empírico, el algoritmo que exponemos aquí utiliza un método de aproximación estocástica. En realidad el método propuesto a continuación es una versión modificada del AGL original y se conoce como *algoritmo de aprendizaje competitivo* en la literatura

de redes neuronales. En síntesis los pasos que caracterizan dicho algoritmo son los que se muestran a continuación:

1) *Inicio*: dados los puntos de datos  $\mathbf{x}(k)$ ,  $k = 1, 2, \dots$ , y los centros de salida iniciales  $\mathbf{c}_j(0)$ ,  $j = 1, \dots, m$ .

2) Determinar el centro  $\mathbf{c}_j(k)$  más próximo al punto  $\mathbf{x}(k)$ ,

$$j = \arg \min_j L(\mathbf{x}(k), \mathbf{c}_j(k)) \quad (3.29)$$

utilizando la función de error al cuadrado como función de pérdidas. Esto se simplifica para obtener la regla del vecino más próximo,

$$j = \arg \min_j \|\mathbf{x}(k) - \mathbf{c}_j(k)\| \quad (3.30)$$

El hallazgo del centro más próximo se denomina *competición* (entre centros) en los métodos de redes neuronales.

3) *Actualizar* el centro de salida utilizando las ecuaciones,

$$\begin{aligned} \mathbf{c}_j(k_j + 1) &= \mathbf{c}_j(k_j) - \gamma(k_j) \text{grad } L(\mathbf{x}(k), \mathbf{c}_j(k_j)) \\ k_j &= k_j + 1 \end{aligned} \quad (3.31)$$

Obsérvese que cada centro puede tener su propia razón de aprendizaje, lo que se indica con  $k_j$  en  $\gamma(k_j)$ , con  $j = 1, \dots, m$ . La razón de aprendizaje  $\gamma(k_j)$  debe cumplir las condiciones de aproximación estocástica. Cuando la función de pérdidas es el error al cuadrado, el gradiente se calcula como,

$$\frac{\partial L(\mathbf{x}, \mathbf{c}_j)}{\partial \mathbf{c}_j} = \frac{\partial}{\partial \mathbf{c}_j} \|\mathbf{x} - \mathbf{c}_j\|^2 = 2(\mathbf{x} - \mathbf{c}_j) \quad (3.32)$$

con este gradiente, los centros de salida se actualizan por,

$$\begin{aligned} \mathbf{c}_j(k_j + 1) &= \mathbf{c}_j(k_j) - \gamma(k_j) [\mathbf{x}(k) - \mathbf{c}_j(k_j)] \\ k_j &= k_j + 1 \end{aligned} \quad (3.33)$$

que se conoce comúnmente como *regla de aprendizaje competitivo* en redes neuronales.

---

Un problema común en el AGL es que una mala elección de los centros iniciales desemboca en soluciones óptimas localmente malas. Una aproximación para superar este inconveniente consiste en aplicar el AGL varias veces comenzando con distintas condiciones iniciales y luego eligiendo la mejor solución. Otro método en la misma línea es el conocido como *mecanismo de conciencia* (DeSieno, 1988), que es realmente una modificación del procedimiento descrito por la ecuación (3.30). Cada unidad mantiene una traza del número o frecuencia de las veces que han vencido la competición en el paso 2. Sea  $frec_j(k)$  dicha frecuencia correspondiente a la unidad  $j$  en la iteración  $k$ . Entonces la regla del vecino más próximo (3.30) se modifica a,

$$j = \arg \min_j [\|\mathbf{x}(k) - \mathbf{c}_j(k)\| frec_j(k)] \quad (3.34)$$

El paso 3 no cambia. La nueva medida de distancia en (3.34) fuerza a cada unidad a vencer en promedio el mismo número de veces. En otras palabras, las vencedoras más frecuentes se sienten culpables (toman conciencia) y reducen su tendencia a vencer mediante (3.34).

### 3.2.4.2 Fase de decisión

El objetivo de la fase de decisión consiste en clasificar una nueva muestra  $\mathbf{x}_s$  en alguna de las clases  $w_j$  existentes. Con tal propósito, se recuperan de la BC los centros (pesos)  $\mathbf{c}_j$  almacenados durante la fase de entrenamiento, determinándose la proximidad de la muestra a todos los centros de las clases. La proximidad se establece en función de un criterio de mínima distancia. Esta distancia debe ser la misma que la utilizada en la fase de entrenamiento, por tanto la Euclídea. La clasificación de la muestra se lleva a cabo según la regla de decisión siguiente,

$$\mathbf{x}_s \in w_j \quad \text{si} \quad \|\mathbf{x}_s - \mathbf{c}_j\| < \|\mathbf{x}_s - \mathbf{c}_k\| \quad \forall k \neq j \quad (3.35)$$

Las salidas proporcionadas por este clasificador son las distancias euclídeas entre  $\mathbf{x}_s$  y cada uno de los centros  $\mathbf{c}_j$ , es decir  $d(\mathbf{x}_s, \mathbf{c}_j) = \|\mathbf{x}_s - \mathbf{c}_j\|$ .

### 3.2.5 Mapas Auto-Organizativos

#### 3.2.5.1 Fase de entrenamiento

Es uno de los métodos más populares en redes neuronales utilizado para la reducción de la dimensionalidad de los datos por su capacidad para representar éstos por vectores significativos. Este método se conoce en la terminología inglesa como *Self-Organizing Maps* (SOM). En este trabajo utilizaremos estas siglas por su amplia difusión en la comunidad científica. Como otros muchos modelos en el campo de las redes neuronales fue propuesto originalmente como una aplicación inspirada en un fenómeno biológico por Kohonen (1989, 1990, 1997). Su descripción detallada se puede encontrar en Pajares y Cruz (2007) o Pandya y Macy (1996).

El modelo de Kohonen posee una arquitectura que consta de dos capas, una capa de entrada y una capa de Kohonen, también denominada mapa o capa de salida. Dichas capas se encuentran totalmente inter-conectadas. Cada neurona en la capa de entrada tiene una conexión hacia cada neurona en la capa de salida. En la figura 3.3 se ilustra la arquitectura de este tipo de redes en el caso bi-dimensional.

Cada una de las conexiones entre una neurona  $i$  de la capa de entrada y una neurona  $j$  de la capa de salida tiene asociado un peso, denominado peso de conexión  $c_{ij}$ , de forma que la componente  $x_i$  del vector de entrada  $\mathbf{x} \in \mathbb{R}^q$  es inyectada a través de  $i$ , que está conectada a  $j$  mediante la conexión y su peso. Esto significa que cada neurona  $j$  de la capa de salida tiene asociado un vector de pesos  $\mathbf{c}_j \equiv \{c_{1j}, \dots, c_{ij}, \dots, c_{qj}\}$  cuyas componentes son los pesos de conexión con las neuronas de la capa de entrada. Estos pesos son los que se actualizan durante el proceso de entrenamiento, siendo el objetivo del aprendizaje.

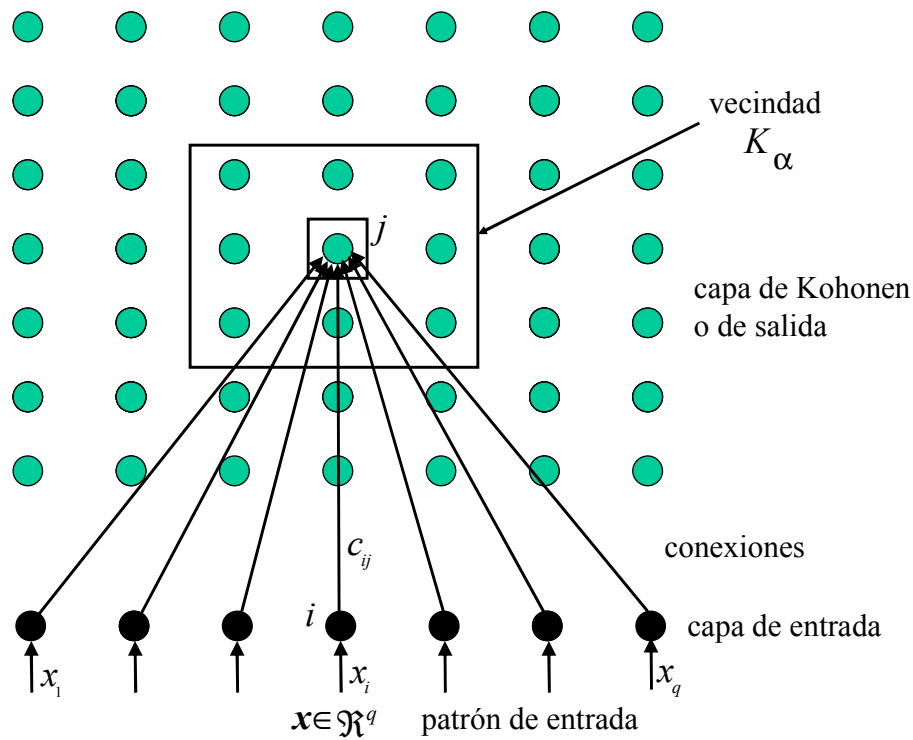


Figura 3.3 Modelo de arquitectura de Kohonen en el caso bi-dimensional

Dada una muestra de entrada  $\mathbf{x}$ , el método determina la neurona vencedora de la capa de salida. El concepto de neurona vencedora se asocia, en este caso, con el vector de pesos  $\mathbf{c}_j$  que mejor se aproxime a la muestra de entrada.

Desde el punto de vista de las redes competitivas, se diría que es la neurona vencedora la que actualiza los pesos de conexión. No obstante, el procedimiento introducido en este algoritmo no considera en exclusividad neuronas vencedoras sino que dependiendo del grado de proximidad del patrón de entrada a las distintas neuronas, así será el grado de actualización de los correspondientes pesos de conexión.

El concepto de proximidad se establece mediante la *función de vecindad* dada por la siguiente expresión,

$$K_{\alpha(t)}(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\alpha^2(t)}\right) \quad \text{con} \quad \alpha(t) = \alpha_{\text{inicial}} \left(\frac{\alpha_{\text{final}}}{\alpha_{\text{inicial}}}\right)^{t/t_{\text{max}}} \quad (3.36)$$

donde  $t$  es el número de iteración y  $t_{\text{max}}$  es el número máximo de iteraciones, que se debe especificar de antemano. El ancho inicial de la vecindad  $\alpha_{\text{inicial}}$  se elige de modo que la



propia vecindad cubra todas las neuronas y el ancho final de la vecindad  $\alpha_{final}$  controla el grado de variación de la función de vecindad entre iteraciones consecutivas. Estos parámetros permanecen invariables una vez fijados. Se dice que  $\mathbf{x}$  está en la vecindad de  $\mathbf{x}'$  según el valor que tome dicha función, de suerte que fijado un determinado umbral, los vectores con valores por debajo del mismo se consideran que no pertenecen a la región de vecindad y viceversa. No obstante, en este trabajo no se aplica este criterio; por tanto todas las muestras mantienen un cierto grado de pertenencia a una vecindad dada, por pequeño que éste sea y por consiguiente no es necesario fijar dicho umbral.

El algoritmo de auto-organización se sintetiza de acuerdo con los pasos siguientes (Pandya y Macy, 1996),

1. *Inicialización*: inicializar el vector de pesos  $\mathbf{c}_j$  ( $t = 0$ ), que bien pudiera ser a valores aleatorios o conocidos, siempre y cuando en este último caso, se hayan obtenido mediante algún otro procedimiento. Inicializar la razón de aprendizaje  $\gamma(t = 0)$ . Definir los valores involucrados en  $\alpha(t)$ , a saber:  $\alpha_{inicial}$ ,  $\alpha_{final}$  y  $t_{max}$ .
2. Para cada muestra representada por su vector  $\mathbf{x}$  realizar los pasos 3 y 4.
3. *Medida de similitud*: seleccionar la neurona  $j$  cuyo vector de pesos esté más próximo al patrón de entrada según el criterio definido en (3.37), donde  $\|\cdot\|$  define una norma, elegida como la Euclídea en este trabajo,

$$j \text{ vence si } \|\mathbf{x} - \mathbf{c}_j\| < \|\mathbf{x} - \mathbf{c}_k\| \quad \forall k \neq j \quad (3.37)$$

4. *Entrenamiento*: actualizar los pesos de las neuronas dentro de la región de vecindad de acuerdo con la siguiente ley,

$$\mathbf{c}_j(t+1) = \begin{cases} \mathbf{c}_j(t) + \gamma(t)[\mathbf{x} - \mathbf{c}_j(t)] & j \in K_{\alpha(t)} \\ \mathbf{c}_j(t) & \text{de cualquier otro modo} \end{cases} \quad (3.38)$$

5. Actualizar la razón de aprendizaje, según la ley definida con tal propósito, como se verá posteriormente.
6. Disminuir el valor de la función de vecindad mediante la variación de  $\alpha(t)$ .

7. Verificar la condición de convergencia a partir del criterio definido en (3.39), según el cual cuando la variación de todos los pesos entre dos iteraciones consecutivas está por debajo de un determinado umbral  $\varepsilon$ , según la norma Euclídea, el proceso se detiene,

$$\|\mathbf{c}_j(t+1) - \mathbf{c}_j(t)\| < \varepsilon \quad \forall j \quad (3.39)$$

La ley de actualización de pesos tanto para las neuronas vencedoras como para las que no han vencido se expresa como sigue,

$$\mathbf{c}_j(t+1) = \mathbf{c}_j(t) + \gamma(t) K_{\alpha(t)}(\mathbf{x}, \mathbf{c}_j(t)) [\mathbf{x} - \mathbf{c}_j(t)] \quad (3.40)$$

Obsérvese cómo la función de vecindad y la razón de aprendizaje determinan conjuntamente el valor de la fuerza ejercida por  $\mathbf{x}$  sobre  $\mathbf{c}_j$  para la actualización de éste.

Los valores de los parámetros involucrados en este algoritmo se determinan, como en el caso de los anteriores, mediante validación cruzada como se verá en el capítulo cuatro, sección 4.2.1. Finalmente, la razón de aprendizaje se ha establecido como  $\gamma(t) = 1/(10+t)$ , que determina un decremento de su valor a medida que el número de iteraciones aumenta; lo cual está en concordancia con el hecho de que a mayor experiencia en la vida menor es el nivel de aprendizaje obtenido.

La aplicación del algoritmo SOM al caso de la clasificación de las texturas de las imágenes naturales requiere la definición de los patrones de entrada así como de los pesos de conexión y el número de neuronas en la capa de Kohonen. De este modo, los vectores de entrada son los vectores  $\mathbf{x} \in \mathbb{R}^3$  que caracterizan las componentes espectrales de los píxeles según el modelo de color RGB adoptado en nuestro caso. Por tanto, el número de neuronas en la capa de entrada es tres, que se corresponden con cada una de las tres componentes R, G y B utilizadas.

Por otra parte, el número de neuronas en la capa de salida viene determinado por el número de clases existentes, donde el vector de pesos de conexión de cada neurona representa el centro de cada clase, es decir  $\mathbf{c}_j$  para la clase  $w_j$ . Estos vectores de pesos o centros son los parámetros de aprendizaje que se almacenan en la BC tras el proceso de entrenamiento, según el esquema de la figura 3.1.

### 3.2.5.2 Fase de decisión

Nuevamente, el objetivo de la fase de decisión consiste en clasificar una nueva muestra  $\mathbf{x}_s$  en alguna de las clases  $w_j$  existentes. Con tal propósito, se recuperan de la BC los centros (pesos)  $\mathbf{c}_j$  almacenados durante la fase de entrenamiento, determinándose la proximidad de la muestra a todos los centros de las clases. La proximidad, en este caso, se establece también en función de un criterio de mínima distancia. Esta distancia debe ser la misma que la utilizada en la fase de entrenamiento, por tanto la Euclídea. La clasificación de la muestra se lleva a cabo según la regla de decisión mostrada en la ecuación (3.35).

Las salidas proporcionadas por este clasificador son las distancias euclídeas entre  $\mathbf{x}_s$  y cada uno de los centros  $\mathbf{c}_j$ , es decir,  $d(\mathbf{x}_s, \mathbf{c}_j) = \|\mathbf{x}_s - \mathbf{c}_j\|$ .

## 3.3 Funciones criterio para la partición

Para estimar los parámetros derivados del proceso de aprendizaje, los cinco métodos expuestos en la sección 3.2 necesitan la distribución en  $c$  clases,  $w_1, w_2, \dots, w_c$  de las  $n$  muestras contenidas en  $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\} \in \mathcal{R}^q$ ; esto es, la partición de dichas muestras. Dicha distribución puede realizarse de dos formas, una es manualmente bajo la supervisión de un experto; la otra es de forma automática sin necesidad de supervisión. Esta última es la adoptada en este trabajo de investigación con el fin último de conseguir la automatización del proceso de aprendizaje, que constituye uno de los objetivos de la investigación.

Se plantea ahora el problema de hasta qué punto la partición de las muestras sin supervisión puede considerarse como válida. En esta sección se aborda precisamente esta problemática. Con tal propósito, en primer lugar se establecen las premisas previas, que determinan cuándo una partición se considera válida y en segundo lugar, cómo se mide el grado de validez de dicha partición.

Resulta comúnmente aceptado que cuanto más similares son entre sí las muestras de una misma clase y a la vez más diferencias existen entre las muestras de clases distintas, tanto mejor será la partición. Por tanto, el objetivo consiste en obtener una partición que cumpla lo mejor posible las dos premisas anteriores.

La validación de las particiones ha sido un tema ampliamente tratado en la literatura, como puede deducirse de los estudios realizados por Volkovich y col. (2008) y referencias asociadas. A continuación abordamos esta cuestión, teniendo en cuenta la vinculación de los métodos de validación con los clasificadores utilizados.

Como veremos en el punto 3.4, el clasificador en el que nos basamos para realizar la partición inicial en clases es el Agrupamiento Borroso, que estima los grados de pertenencia de las muestras a las clases mediante la ecuación (3.3). Con este fin se han considerado por lo tanto, diferentes funciones criterio que consisten en medidas escalares para validar dicha partición inicial, que son el *coeficiente de partición* (PC), la *entropía de la partición* (PE) y el *índice de Xie-Beni* (XB).

Por otro lado, existen aún dos cuestiones útiles relacionadas con las funciones criterio que pueden ser aprovechadas convenientemente. En primer lugar, como se verá en la sección 3.4 durante el estudio del clasificador combinado, uno de los problemas que surgen consiste en determinar cuál es la importancia relativa de cada clasificador individual dentro de la combinación, es decir cuál de ellos presenta un mejor o peor comportamiento. En este trabajo se ha diseñado un procedimiento para deducir el comportamiento de los diferentes clasificadores individuales en función de los valores proporcionados por las funciones criterio. En segundo lugar, los clasificadores individuales estudiados hasta ahora, excepto el clasificador paramétrico Bayesiano, son de naturaleza supervisada, bien por su propia definición o por aproximación. Pues bien, gracias al comportamiento de las funciones criterio, es posible automatizar el proceso y conseguir un diseño no supervisado.

A continuación, se definen cada una de las funciones criterio para la validación de las clases, cuya base es el propio clasificador borroso. En sección 3.3.1 se analiza el comportamiento de estas funciones según el número de clases.

#### 1. *Coeficiente de partición* (PC):

El *coeficiente de partición* (PC) es una función criterio que mide la cantidad de solapamiento entre las clases, y se define como sigue (Zimmermann, 1991; Kim y col., 2003):

$$PC(U; c) = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^c (\mu_{ij})^2 \quad \text{con} \quad \frac{1}{c} \leq PC(U; c) \leq 1 \quad (3.41)$$

donde  $U$  es la matriz de grados de pertenencia  $\mu_{ij}$  a las clases  $w_j$ , definida en la sección 3.2.1.1;  $n$  es el número de muestras y  $c$  el número de clases. Esta nomenclatura es común para las definiciones del resto de funciones criterio.

Obsérvese que el dominio de  $PC$  se reduce al intervalo  $[1/c, 1]$ . Según este coeficiente, las mejores particiones se obtienen para valores máximos de  $PC$ , como se analizará posteriormente.

### 2. Coeficiente de entropía (PE):

Al igual que  $PC$ , mide la cantidad de solapamiento entre las clases; se define como sigue (Kim y col., 2003):

$$PE(U; c) = -\frac{1}{n} \sum_{j=1}^c \sum_{i=1}^n [\mu_{ij} \log(\mu_{ij})] \quad \text{con} \quad 0 \leq PE(U; c) \leq \log(c) \quad (3.42)$$

Obsérvese que el dominio de  $PE$  se reduce al intervalo  $[0, \log(c)]$ . Según este coeficiente, las mejores particiones se obtienen para valores mínimos de  $PE$ , como se analizará posteriormente.

### 3. Índice de Xie-Beni (XB):

Xie Beni propone un índice de validez que se centra en dos propiedades, la compactación y la separación, y se define con la siguiente ecuación (Kim y col., 2003):

$$XB(U; c) = \frac{\sum_{i=1}^c \sum_{j=1}^n \mu_{ij}^2 \|x_j - v_i\|^2}{n \left( \min_{i \neq k} \|v_i - v_k\|^2 \right)} \quad (3.43)$$

Donde el numerador indica la compactación de la partición fuzzy y el denominador la fuerza de la separación entre las clases. Según este coeficiente, las mejores particiones se obtienen minimizando  $XB$  sobre  $c = 2, 3, \dots, c_{\max}$ .

Existen otras funciones basadas en la compactación y la separación de clases, tales como las de *Fukuyama y Sugeno* o *Kwon* entre otras, pudiéndose encontrar en Kim y col., (2003); si bien tras los diversos experimentos realizados, su comportamiento no introduce ninguna aportación relevante con respecto a las citadas anteriormente, motivo por el cual no se han considerado.

### 3.3.1 Comportamiento de las funciones criterio

La función criterio  $PC$  es monótona creciente, mientras que  $PE$  y  $XB$  son monótonas decrecientes, en los tres casos en relación al número de clases; llegando a alcanzar sus valores extremos exactamente cuando el número de muestras es igual al número de clases ( $c = n$ ). En el caso de  $PC$  dicho valor extremo es la unidad, que coincide con el máximo; mientras que para  $PE$  y  $XB$  los extremos son los valores mínimos, que se corresponden con el cero. Obviamente la obtención de una partición con un número de clases igual al número de muestras, donde cada clase posee una única muestra, no tiene ningún sentido a efectos prácticos. Por este motivo, el objetivo se centra ahora en obtener un criterio para determinar bajo qué condiciones de máximo o mínimo se puede considerar válida la partición.

Dada la complejidad de las imágenes forestales tratadas, en ocasiones es muy complicado determinar exactamente el número total de texturas diferentes. Si se configura un criterio de validación muy estricto y se generan demasiadas clases (en nuestro caso más de seis), es muy complicado determinar cuáles de ellas cubren totalmente las texturas existentes en las imágenes y en especial las relacionadas con los troncos de los árboles, que constituyen las texturas de interés prioritario.

Es por esto que para este tipo de imágenes, no conviene que los criterios de partición sean muy estrictos, ni que las funciones criterio tengan demasiada variación, con el fin de poder jugar con ellos y finalmente determinar cuál es el número de clases necesario para una mejor clasificación de los troncos.

En las pruebas realizadas detalladas en el apartado 4.3.2, se han demostrado más estables los índices  $PC$  y  $PE$ , mientras que  $XB$  varía demasiado dependiendo del tipo de imágenes y de los píxeles elegidos para formar la partición inicial.

Por otro lado, en lugar de tratar de determinar qué función criterio se comporta mejor que las otras a la hora de validar la partición inicial generada, se ha tratado de combinar los criterios más estables, es decir, PC y PE, hallando valores umbral de manera que cuando la variación relativa se encuentre por debajo de los mismos, la partición se considera válida. Los detalles concretos sobre la obtención de este umbral se posponen para la sección 4.3.2.

### 3.4 Método de clasificación combinado: MCDMF

En la sección anterior se han descrito cinco métodos de clasificación catalogados como individuales, que constituyen la base del método combinado propuesto en este trabajo. Dichos clasificadores son los que se mencionan a continuación, identificados a partir de ahora con la simbología que se indica: *a)* el método de Agrupamiento Borroso o “Fuzzy Clustering” (*FC*), *b)* el estimador paramétrico Bayesiano (*PB*), *c)* el estimador no paramétrico basado en la ventana de Parzen (*PZ*), *d)* el algoritmo generalizado de Lloyd (*AGL*), y *e)* el método basado en los Mapas Auto-Organizativos o “Self-Organizing Maps” (*SOM*).

En Duda y col. (2001) se plantea el diseño de un clasificador *no supervisado* a partir de otro *supervisado* durante la fase de entrenamiento. Básicamente, el procedimiento consiste en un proceso iterativo que comienza suponiendo la existencia de un cierto número de clases, de suerte que dado un clasificador, éste estima sus parámetros bajo dicha suposición. Posteriormente, se evalúa la calidad de la partición utilizando una serie de medidas. Si las medidas realizadas son satisfactorias según el criterio o criterios establecidos, la partición se considera válida. Este es el esquema seguido en el diseño del método combinado no supervisado propuesto, como aportación fundamental en este trabajo en el contexto de las imágenes naturales de exterior, como se verá posteriormente. Este método realiza la combinación a nivel de píxel, catalogada por este motivo como de naturaleza *local*. Las características son, por tanto, los píxeles. Las tres componentes espectrales de estas características en el modelo de color *RGB*, son las propiedades utilizadas, tal y como se ha descrito en la sección 3.1.1. La combinación de los clasificadores individuales se realiza durante la fase de decisión, abordándose al igual que en los métodos individuales, desde las dos fases típicas involucradas en todo proceso de aprendizaje/clasificación, a saber: *entrenamiento* y *decisión*, figura 3.1.

Bajo la cobertura del enfoque local se propone un método que utiliza los cinco clasificadores individuales citados previamente, esto es: FC, PB, PZ, AGL, y SOM. El método se basa en el paradigma de la *teoría de la Toma de Decisiones Multicriterio*, conocida en la literatura especializada como “*Multi-Criteria Decision Making, MCDM*”, bajo la perspectiva de la lógica *Fuzzy*. A este método nos referiremos de aquí en adelante con las siglas bajo la nomenclatura internacional *MCDMF*, por su amplia difusión en la comunidad científica internacional.

Resumiendo, en esta sección se propone una estrategia local a nivel de píxel para la combinación de clasificadores durante la fase de decisión, mediante el método MCDMF. A continuación se describe dicho método combinado, dando detalles del mismo tanto en lo que respecta a la fase de entrenamiento como de decisión.

La figura 3.4 muestra el esquema de diseño del clasificador no supervisado basado en el paradigma MCDMF. Como puede observarse, el procedimiento trabaja en las ya mencionadas dos fases de entrenamiento y decisión. En esta sección se expone brevemente el mecanismo de funcionamiento durante ambas fases, poniendo el énfasis en la de decisión, ya que es en ella donde realmente se produce el aporte fundamental de este trabajo de investigación en relación a este clasificador combinado. Con tal propósito, en primer lugar se describe el proceso de entrenamiento donde intervienen los cinco clasificadores individuales, que estiman los parámetros involucrados en los respectivos procesos. Posteriormente, se exponen las líneas generales del método MCDMF sobre el que se sustenta el procedimiento de combinación propuesto. Finalmente, se concreta la aplicación del método al problema de la clasificación de las texturas naturales, que constituye el objeto de la investigación.

La explicación de algunos detalles mostrados en la figura 3.4 tales como los relativos a criterios de beneficio y coste se posponen para su posterior explicación en las siguientes secciones.



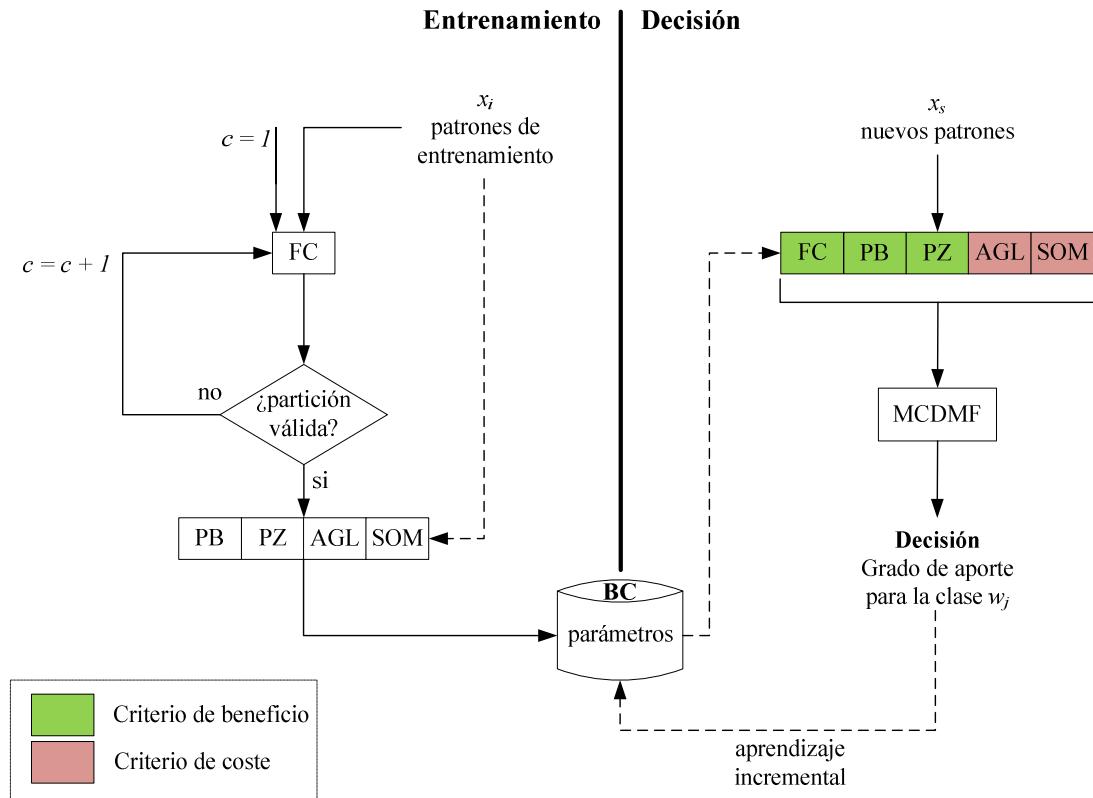


Figura 3.4 Diseño de un clasificador no supervisado combinado:  
Fases de entrenamiento y decisión.

### 3.4.1 Fase de entrenamiento

El proceso de entrenamiento comienza con el procesamiento de los patrones o muestras de entrenamiento disponibles en cada momento, constituyendo las entradas al sistema. Inicialmente, se supone la existencia de una única clase y que todas las muestras pertenecen a dicha clase, por ello  $c = 1$ . Bajo este supuesto, se desencadena el procedimiento de entrenamiento siguiendo el proceso descrito en la sección 3.2.1.1; llegándose a establecer una partición de las muestras en la única clase existente en este momento. La partición con una única clase se considera por definición no válida, ya que en las imágenes disponibles esta situación no se produce nunca. Por esta razón, se intenta una nueva partición con  $c = 2$ . Tras lo cual, se valora si la nueva partición es o no válida.

El proceso de validación se lleva a cabo mediante la combinación de los criterios PC y PE, tal como explicábamos en el punto 3.3.1. Si la partición no es válida, según el criterio anterior, el número de clases  $c$  se incrementa en una unidad, procediéndose de nuevo a repetir el proceso previo hasta conseguir la validación de la partición. La distribución de

las muestras en las clases, una vez conocido el número de éstas, se lleva a cabo mediante el proceso pseudo-aleatorio descrito en la sección 3.2.1.1 según la teoría expuesta en Balasko y col. (2006). Este es el proceso básico que confiere a este diseño su naturaleza no supervisada según la filosofía establecida en Duda y col. (2001), como se ha mencionado previamente.

Por el contrario, cuando la partición es considerada válida, también según el mismo criterio anterior, los cinco clasificadores individuales realizan sus correspondientes procesos de entrenamiento para llevar a cabo la estimación de sus parámetros. Estos procesos de entrenamiento son los descritos en las secciones que se indican a continuación: FC en 3.2.1.1, PB en 3.2.2.1, PZ en 3.2.3.1, AGL en 3.2.4.1 y SOM en 3.2.5.1.

Tras los respectivos procesos de entrenamiento, los parámetros estimados o aprendidos por cada uno de esos clasificadores son almacenados en la BC, quedando disponibles para su posterior recuperación y uso durante la fase de decisión, donde se produce la combinación de los clasificadores, figuras 3.1 y 3.4.

## **3.4.2 Descripción general del clasificador combinado**

### **3.4.2.1 Formulación del problema y definiciones**

La toma de decisiones es el proceso por el cual se intenta encontrar la mejor opción, entre un conjunto de alternativas posibles, según una serie de criterios disponibles. En la mayoría de los problemas de decisión se plantea la cuestión de resolver un problema con múltiples criterios. Un problema de esta naturaleza puede expresarse en forma matricial según se expresa en la ecuación (3.44). Para ello se dispone de  $o$  alternativas,  $A_j (j=1,2,...,o)$  y  $n$  criterios  $CR_i (i=1,2,...,n)$ . Las alternativas determinan las opciones que el clasificador puede elegir a la hora de tomar la decisión. Los criterios representan la base sobre los que dicha decisión va a ser tomada. Cada criterio  $CR_i$ , tiene asociado un peso específico  $z_i$  que determina la importancia o valor concedido al mismo. Con ello se definen las dos estructuras siguientes (Wang y Fenton, 2006): a)  $D$ , conocida como la *matriz de decisión*, en ella cada entrada  $x_{ji}$  representa la valoración de la

alternativa  $A_j$  con respecto al criterio  $CR_i$  y b)  $Z$ , que representa el *vector de pesos* específicos asociados a cada criterio.

$$D = \begin{matrix} & CR_1 & CR_2 & \cdots & CR_n \\ \begin{matrix} A_1 \\ A_2 \\ \vdots \\ A_o \end{matrix} & \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{o1} & x_{o2} & \cdots & x_{on} \end{bmatrix} \end{matrix} \text{ y } Z = [z_1, z_2, \dots, z_n], \text{ donde } i=1, 2, \dots, n \text{ y } j=1, 2, \dots, o \quad (3.44)$$

En forma abreviada, la ecuación (3.44) puede expresarse como sigue,

$$D = [x_{ji}]_{o \times n} \text{ y } Z = [z_i]_{1 \times n}, \text{ donde } i=1, 2, \dots, n \text{ y } j=1, 2, \dots, o \quad (3.45)$$

Los criterios se enmarcan según las dos categorías siguientes:

- *criterio de beneficio*: donde el valor más alto de  $x_{ji}$  representa la mejor opción a la hora de tomar la decisión con respecto a dicho criterio.
- *criterio de coste*: donde el valor más bajo de  $x_{ji}$  representa la mejor opción para la toma de la decisión.

En los métodos clásicos sobre toma de decisiones, los valores numéricos tanto de los criterios como de los pesos se conocen de forma precisa. Una revisión exhaustiva sobre tales métodos se puede encontrar en los trabajos de Hwang y Yoon (1981). Una de las primeras propuestas formales sobre la toma de decisiones se encuentra en dicho trabajo, siendo conocida como TOPSIS ("*Technique for Order Performance by Similarity to Ideal Solution*"), que basa la elección según el concepto de que la alternativa elegida será la que obtenga la mínima distancia a una solución ideal positiva, a la vez que se obtiene la máxima distancia a una solución ideal negativa.

En algunas circunstancias no claramente identificables ni identificadas, los datos precisos son inapropiados en algunos problemas de la vida real. Esto es debido a que los juicios emitidos, desde el punto de vista humano, sobre determinadas preferencias son a menudo vagos o imprecisos, no resultando fácil estimar su preferencia mediante un valor numérico exacto.

Una posible solución al modelo de valores exactos, proviene del ámbito de la lógica fuzzy. Tal es el caso de los primeros trabajos propuestos en Zadeh (1975), que utilizan etiquetas lingüísticas. Teniendo en cuenta este enfoque, tanto los criterios como las alternativas toman valores con un cierto grado de incertidumbre. Bajo esta perspectiva se redefine la ecuación (3.45) como sigue,

$$\tilde{D} = [\tilde{x}_{ji}]_{o \times n} \text{ y } \tilde{Z} = [\tilde{z}_i]_{1 \times n}, \text{ donde } i=1,2,\dots,n \text{ y } j=1,2,\dots,o \quad (3.46)$$

donde ahora  $\tilde{x}_{ji}$  representa la valoración fuzzy de la alternativa  $A_j$  con respecto al criterio  $CR_i$ ; y  $\tilde{z}_i$  determina el peso específico o importancia relativa fuzzy concedida al correspondiente criterio  $CR_i$ . En particular, un enfoque fácil y eficaz para capturar la incertidumbre por parte de un experto sobre el valor de un número desconocido, se puede obtener como un *número triangular fuzzy*. A partir de estos conceptos se definen del mismo modo las soluciones ideales tanto positiva como negativa fuzzy, por tanto, cada uno de los valores  $\tilde{x}_{ji}$  y  $\tilde{z}_i$  definidos en sendas matrices de la ecuación (3.46) y representando etiquetas lingüísticas, pueden describirse por números triangulares fuzzy.

*Definición:* un número triangular fuzzy  $\tilde{a}$  se puede definir por una terna  $(a_1, a_2, a_3)$ . La función de pertenencia de un valor  $x$  dado, se define por medio de la ecuación (3.47) según Kaufmann y Gupta (1985),

$$\mu_{\tilde{a}}(x) = \begin{cases} (x - a_1)/(a_2 - a_1), & a_1 \leq x \leq a_2 \\ (a_3 - x)/(a_3 - a_2), & a_2 \leq x \leq a_3 \\ 0, & \text{en cualquier otro caso} \end{cases} \quad (3.47)$$

Los números triangulares fuzzy están basados en la evaluación de tres valores: el mínimo/máximo valor posible  $a_1/a_3$  el valor más probable  $a_2$  y el máximo/mínimo valor posible  $a_3/a_1$ . En Chen (2000) se puede encontrar un estudio pormenorizado sobre las definiciones y propiedades de estos números fuzzy y las variables lingüísticas asociadas. De este modo, las variables lingüísticas que aparecen en la ecuación (3.46) pueden describirse específicamente mediante números triangulares fuzzy de la forma,  $\tilde{x}_{ji} = (x_{ji1}, x_{ji2}, x_{ji3})$  y  $\tilde{z}_i = (z_{i1}, z_{i2}, z_{i3})$ .

### 3.4.2.2 Normalización

Los valores de los criterios pueden presentarse según diferentes escalas de valores. Esto es lo que ocurre en nuestra aplicación dado que las decisiones se basan en las salidas proporcionadas por los clasificadores, que representarán los criterios de decisión. Dichas salidas son por un lado grados de pertenencia (FC) o probabilidades (PB, PZ) que varían en el rango  $[0,1]$  y por otro distancias (AGL, SOM), donde ahora los rangos de valores difieren de los anteriores. Por este motivo, para que durante la toma de decisiones el hecho de disponer de datos en diferentes rangos de valores no suponga un obstáculo, es necesario proceder a realizar un proceso de normalización.

Específicamente, se trata de normalizar los números fuzzy de la matriz de decisión, de forma que a partir de esta normalización se genere lo que denominamos *matriz de decisión normalizada*, también conocida en la terminología inglesa como “*performance matrix*”. Dicha matriz queda identificada en función de sus elementos como sigue (Wang y Fenton, 2006; Gu y Zhu, 2006),

$$\bar{P} = [\tilde{P}_{ji}]_{o \times n} \quad (3.48)$$

Para evitar fórmulas complicadas de normalización, tal como las utilizadas en TOPSIS por Chen (2000), se propone una transformación de escalado lineal. Esta es la adoptada en este trabajo de investigación, la cual permite transformar los diferentes valores de los criterios en valores comparables entre sí. Dicha transformación se concreta como sigue,

$$\tilde{p}_{ij} = \begin{cases} \left( \frac{x_{ji1}}{M}, \frac{x_{ji2}}{M}, \frac{x_{ji3}}{M} \right), & M = \max_j x_{ji3}, \text{ } CR_i \text{ criterio de beneficio} \\ \left( \frac{m}{x_{ji3}}, \frac{m}{x_{ji2}}, \frac{m}{x_{ji1}} \right), & m = \min_j x_{ji1}, \text{ } CR_i \text{ criterio de coste} \end{cases} \quad (3.49)$$

Mediante este proceso de normalización el rango de valores de los números fuzzy triangulares se restringen al intervalo  $[0, 1]$ .

### 3.4.2.3 Proceso de decisión

Considerando la importancia de cada criterio, se puede construir la matriz de decisión fuzzy promediada y normalizada como sigue,

$$\tilde{V} = [\tilde{v}_{ji}]_{o \times n} \text{ donde } i=1,2,\dots,n \text{ y } j=1,2,\dots,o \text{ y } \tilde{v}_{ji} = \tilde{p}_{ji}(\cdot)\tilde{z}_i \quad (3.50)$$

Los elementos de la matriz anterior siguen siendo números fuzzy triangulares positivos y normalizados en el rango  $[0,1]$ . Las soluciones ideales positiva y negativa vienen dadas por la siguiente expresión,

$$A^+ = (\tilde{v}_1^+, \tilde{v}_2^+, \dots, \tilde{v}_n^+) \text{ y } A^- = (\tilde{v}_1^-, \tilde{v}_2^-, \dots, \tilde{v}_n^-) \quad (3.51)$$

donde  $\tilde{v}_i^+ = (1,1,1)$  y  $\tilde{v}_i^- = (0,0,0)$  con  $i=1,2,\dots,n$ .

La distancia de cada alternativa con respecto a  $A^+$  y  $A^-$  se calcula como,

$$d_j^+ = \sum_{i=1}^n d(\tilde{v}_{ji}, \tilde{v}_i^+) \text{ y } d_j^- = \sum_{i=1}^n d(\tilde{v}_{ji}, \tilde{v}_i^-) \quad j=1,2,\dots,o \quad (3.52)$$

donde  $d(\cdot, \cdot)$  es la distancia entre dos números fuzzy. El *método del vértice*, definido en Chen (2000), permite calcular la *distancia entre dos números fuzzy* según la ecuación (3.53); dados  $\tilde{a} = (a_1, a_2, a_3)$  y  $\tilde{b} = (b_1, b_2, b_3)$  dos números triangulares fuzzy positivos, su distancia resulta ser,

$$d(\tilde{a}, \tilde{b}) = \sqrt{\frac{1}{3}[(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2]} \quad (3.53)$$

El objetivo final del problema de decisión consiste en seleccionar una alternativa, con tal propósito, se define el *coeficiente de proximidad* a partir de las distancias definidas en la ecuación (3.52) obtenidas para cada alternativa  $A_j$  ( $j = 1, 2, \dots, o$ ) con respecto a las respectivas distancias ideal positiva y negativa. Dicho coeficiente viene definido en Chen (2000) como,

$$cp_j = \frac{d_j^-}{d_j^+ + d_j^-} \quad \text{con } j=1,2,\dots,o \quad (3.54)$$

Suponiendo que una alternativa real coincide exactamente con el valor ideal positivo, entonces  $d_j^+$  es cero (distancia mínima) y  $d_j^-$  es la unidad (distancia máxima), en cuyo caso  $cp_j$  es igual a la unidad. En el otro extremo, cuando una alternativa toma el valor

ideal negativo ocurre lo contrario, es decir  $d_j^-$  es cero y  $d_j^+$  es la unidad, en cuyo caso  $cp_j$  toma el valor nulo. Esto implica que cuanto más se aproxime  $cp_j$  a uno, más próximo estará  $A_j$  de la solución ideal positiva y más alejado de la negativa. Por tanto, de acuerdo con el coeficiente de proximidad, se puede determinar el orden de preferencia de las alternativas y seleccionar la mejor de entre todas ellas según el criterio del máximo valor de dicho coeficiente.

En resumen, el algoritmo de decisión para el caso multi-alternativa y multi-criterio para la toma de decisiones en el contexto fuzzy se concreta en los pasos siguientes:

- 1) Identificar los criterios  $CR_i$  para la elección de alternativas  $A_j$ .
- 2) Elegir las variables lingüísticas apropiadas.
- 3) Obtener las valoraciones fuzzy  $\tilde{x}_{ji}$  de cada alternativa  $A_j$  con respecto al criterio  $CR_i$ ; y determinar los pesos específicos o importancias relativas  $\tilde{z}_i$  concedidas a cada uno de los criterios.
- 4) Construir la matriz de decisión fuzzy  $\tilde{D}$  y la matriz fuzzy  $\tilde{P}$  normalizada.
- 5) Construir la matriz de decisión fuzzy promediada normalizada con los pesos asociados a cada criterio,  $\tilde{V}$
- 6) Determinar las soluciones ideales positiva y negativa fuzzy,  $A^+$  y  $A^-$ .
- 7) Calcular la distancia de cada alternativa a las soluciones ideales positiva y negativa,  $d_j^+$  y  $d_j^-$  respectivamente.
- 8) Calcular el coeficiente de proximidad de cada alternativa,  $cp_j$ .
- 9) Ordenar las alternativas según el coeficiente de proximidad y seleccionar la mejor alternativa como aquella con el máximo valor de dicho coeficiente.

### 3.4.3 Fase de decisión: combinación de clasificadores

Refiriéndonos de nuevo a la figura 3.4, ante la llegada de una nueva muestra o patrón  $\mathbf{x}_s$ , el problema que se plantea ahora consiste en tomar una decisión sobre la clasificación de dicha muestra en cada una de las clases disponibles y establecidas previamente durante la fase de entrenamiento.

Como bien es sabido, cada clasificador individual toma su propia decisión sobre la clasificación de dicha muestra según las salidas obtenidas en los correspondientes procesos de decisión descritos en la sección 3.2, que sintetizamos de nuevo aquí por simplicidad. Así, FC proporciona los grados de pertenencia de dicha muestra a cada una de las clases; tanto BP como PZ generan probabilidades de pertenencia de la muestra a cada una de las clases; AGL y SOM proporcionan distancias de la muestra a los centros de las clases. Los tres primeros toman la decisión en función de los valores máximos de sus salidas, mientras que los dos últimos lo hacen en función de los valores mínimos.

Pues bien, la idea que subyace a la hora de aplicar el paradigma MCDMF consiste en tomar todos esos valores de salida (grados de pertenencia, probabilidades y distancias) con el fin de combinarlos siguiendo las directrices de dicho paradigma expuestas en la sección 3.4.2 y en base a los trabajos de Wang y Fenton (2006), Gu y Zhu (2006) y Chen (2000). La adaptación del proceso allí descrito se expone a continuación.

Partimos del hecho conocido de que el número de clases ha sido estimado durante la fase de entrenamiento, resultando ser  $\hat{c}$ . Por simplicidad, de aquí en adelante dicho número queda renombrado como  $c$ , por tanto las clases disponibles donde la muestra  $\mathbf{x}_s$  debe ser clasificada son  $w_j$  con  $j = 1, 2, \dots, c$ .

Como se ha mencionado previamente, desde el punto de vista de la teoría de la decisión, el problema consiste en determinar a qué clase pertenece  $\mathbf{x}_s$ , esto supone elegir una de entre varias *alternativas* donde cada clase representa una de tales opciones. Más específicamente, desde el punto de vista de la teoría MCDMF, la elección de una clase es equivalente a elegir una alternativa. Por tanto, a partir de este momento se identifican las clases con las alternativas, esto es  $w_j \equiv A_j$ .

Por otro lado, la elección de una alternativa debe tomarse según ciertos *criterios* previamente establecidos. Como se ha mencionado previamente, los clasificadores individuales toman sus decisiones en función de las salidas obtenidas en la fase de decisión. Por este motivo, los criterios para elegir una alternativa bajo el paradigma MCDMF vienen determinados por las salidas proporcionadas por dichos clasificadores. Siguiendo la teoría expuesta en la sección 3.4.2 existen dos tipos de criterios, a saber: de *beneficio* y de *coste*. En los primeros, los valores mayores representan la mejor opción,



mientras que en los segundos la mejor opción viene dada por los valores más bajos. Pues bien, según lo expuesto anteriormente los clasificadores FC, PB y PZ toman sus decisiones en base a los valores máximos de sus salidas y AGL y SOM lo hacen en función de los mínimos.

En Guijarro (2009), se sugiere considerar los dos grupos de clasificadores por separado, formando en total dos ternas o números triangulares fuzzy. En nuestro caso, como tan sólo disponemos de dos criterios de coste, debemos generalizar este concepto. De este modo permitiremos en un futuro conectar más clasificadores sin alterar el algoritmo global.

Tal como se describe en el apartado 3.4.2.1, un número triangular fuzzy se define como  $(a_1, a_2, a_3)$ , donde  $a_1 / a_3$  es el mínimo/máximo valor posible,  $a_2$  el valor más probable y  $a_3 / a_1$  el máximo/mínimo valor posible. En lugar de formar un número triangular ordenando la salida de tres clasificadores, generamos una terna para cada clasificador, en la que  $a_2$  se corresponde exactamente con la salida del clasificador en cuestión y se generan  $a_1$  y  $a_3$  sumando y restando respectivamente sendos valores aleatorios relativos al rango de valores de cada criterio.

Evidentemente, el cálculo de estos valores aleatorios se limitan, estableciendo el criterio de que nunca superen más del 10% el valor central con el fin de no distorsionar excesivamente los resultados. El cálculo de los límites para los valores aleatorios se ajusta durante la fase experimental mediante el procedimiento de validación cruzada ("*cross-validation*") descrito en el capítulo cuatro, sección 4.2.1 (Duda y col., 2001).

Siguiendo los pasos del procedimiento descrito al final de la sección 3.4.2, a continuación se formaliza la adaptación de nuestro problema al marco del paradigma MCDMF. En lo que sigue nos referiremos a dicho procedimiento de forma genérica siguiendo exactamente esos mismos pasos.

- 1) El primer paso del algoritmo consiste en la *identificación de los criterios*. Pues bien en el caso que nos ocupa los criterios seleccionados son *cada* uno de los clasificadores individuales, según lo explicado anteriormente.

2) El segundo paso consiste en la elección de las *variables lingüísticas*. Atendiendo a los criterios seleccionados, éstas se definen como sigue:

2.1)  $FC$ ,  $PB$  y  $PZ$  toman los valores según los vectores  $FC_s^j \equiv \{\mu_{sj}, P_b(w_j | x_s), P_p(w_j | x_s)\}$ ,  $PB_s^j \equiv \{\mu_{sj}, P_b(w_j | x_s), P_p(w_j | x_s)\}$  y  $PZ_s^j \equiv \{\mu_{sj}, P_b(w_j | x_s), P_p(w_j | x_s)\}$  respectivamente, cuyas componentes son las salidas proporcionadas por los correspondientes clasificadores.

2.2)  $AGL$  y  $SOM$  toman los valores según los vectores:

$$2.2)1. \quad AGL_s^j \equiv \{d(x_s | v_j^0), d(x_s | v_j^1), d(x_s | v_j^2)\}$$

$$2.2)2. \quad SOM_s^j \equiv \{d(x_s | v_j^0), d(x_s | v_j^1), d(x_s | v_j^2)\}$$

Donde  $d(x_s, v_j^r)$  representa la distancia Euclídea entre la muestra  $x_s$  y el centro de cada clase  $w_j$  representados por  $v_j^0$ ,  $v_j^1$  y  $v_j^2$  para  $AGL$  y  $SOM$  respectivamente. Obsérvese, que se ha cambiado la nomenclatura para designar estos centros con respecto a la utilizada para los mismos centros en la sección 3.2.

3) El tercer paso consta de dos partes, la primera consiste en la obtención de las valoraciones fuzzy de cada alternativa con respecto a cada criterio. Esto se lleva a cabo mediante la definición de los siguientes números triangulares fuzzy:

- $FC_s^j \equiv \{FC_{s1}^j, FC_{s2}^j, FC_{s3}^j\}$  con  $FC_{s1}^j \leq FC_{s2}^j \leq FC_{s3}^j$
- $PB_s^j \equiv \{PB_{s1}^j, PB_{s2}^j, PB_{s3}^j\}$  con  $PB_{s1}^j \leq PB_{s2}^j \leq PB_{s3}^j$
- $PZ_s^j \equiv \{PZ_{s1}^j, PZ_{s2}^j, PZ_{s3}^j\}$  con  $PZ_{s1}^j \leq PZ_{s2}^j \leq PZ_{s3}^j$
- $AGL_s^j \equiv \{AGL_{s1}^j, AGL_{s2}^j, AGL_{s3}^j\}$  con  $AGL_{s1}^j \geq AGL_{s2}^j \geq AGL_{s3}^j$
- $SOM_s^j \equiv \{SOM_{s1}^j, SOM_{s2}^j, SOM_{s3}^j\}$  con  $SOM_{s1}^j \geq SOM_{s2}^j \geq SOM_{s3}^j$

La segunda parte del tercer paso implica la obtención de la importancia relativa de cada criterio, que desemboca finalmente en el cómputo de los pesos específicos asociados. Esta tarea se lleva a cabo durante la fase experimental mediante el procedimiento de validación cruzada (“*cross-validation*”) descrito en el capítulo cuatro, sección 4.2.1 (Duda y col., 2001), fijando inicialmente los pesos específicos evaluando los resultados de cada clasificador por separado, utilizando el método descrito en las secciones 4.2.2 y 4.2.3.

- 4) El cuarto paso consiste en la construcción de la matriz de decisión fuzzy y la matriz fuzzy normalizada. Esta tarea se lleva a cabo a partir de la ecuación (3.49). Como se ha mencionado previamente, en nuestro enfoque las alternativas se materializan en la elección de las clases  $w_j$ ; de forma que para llevar a cabo la normalización es preciso obtener los siguientes valores:  $M = \max_j \{X_{s3}^j\}$  y  $m = \min_j \{Y_{s3}^j\}$ , siendo  $X$  relativo a los criterios de beneficio e  $Y$  a los de coste.
- 5) El quinto paso del algoritmo se concreta a partir de las matrices promediadas y normalizadas obtenidas en los pasos previos, cuyos elementos están dados por  $X_{sN}^j$  y  $Y_{sN}^j$ . En la tabla 3.1 se muestra la matriz de decisión normalizada y promediada.
- 6) El sexto paso se refiere a la definición de las soluciones ideales positivas y negativas. En este caso el propio proceso de normalización, tanto para el criterio de beneficio como de coste, restringe directamente los valores al rango  $[0,1]$ , siendo las mejores elecciones las que se toman con valores próximos a 1. Por este motivo, las soluciones ideales positivas y negativas son exactamente las definidas a través de la ecuación (3.51), es decir  $\tilde{v}_i^+ = (1,1,1)$  y  $\tilde{v}_i^- = (0,0,0)$ . En la tabla 3.1 se sintetiza la forma de expresar tanto las alternativas como los criterios, que en realidad viene a ser la materialización de la matriz de decisión definida en la ecuación (3.44), donde las alternativas vienen representadas por las clases y los criterios por los valores de los clasificadores, siendo los datos de decisión los dados por las ternas que representan los números triangulares fuzzy.

		<i>decisión</i>		
		$w_l$	...	$w_c$
<i>criterios</i>  (pesos)	$FC(z_1)$	$FC_{sN}^1 [FC_{s1}^1 / M, FC_{s2}^1 / M, FC_{s3}^1 / M] z_1$		$FC_{sN}^c [FC_{s1}^c / M, FC_{s2}^c / M, FC_{s3}^c / M] z_1$
	$PB(z_2)$	$PB_{sN}^1 [PB_{s1}^1 / M, PB_{s2}^1 / M, PB_{s3}^1 / M] z_2$		$PB_{sN}^c [PB_{s1}^c / M, PB_{s2}^c / M, PB_{s3}^c / M] z_2$
	$PZ(z_3)$	$PZ_{sN}^1 [PZ_{s1}^1 / M, PZ_{s2}^1 / M, PZ_{s3}^1 / M] z_3$		$PZ_{sN}^c [PZ_{s1}^c / M, PZ_{s2}^c / M, PZ_{s3}^c / M] z_3$
	$AGL(z_4)$	$AGL_{sN}^1 [m / AGL_{s1}^1, m / AGL_{s2}^1, m / AGL_{s3}^1] z_4$		$AGL_{sN}^c [m / AGL_{s1}^c, m / AGL_{s2}^c, m / AGL_{s3}^c] z_4$
	$SOM(z_5)$	$SOM_{sN}^1 [m / SOM_{s1}^1, m / SOM_{s2}^1, m / SOM_{s3}^1] z_5$		$SOM_{sN}^c [m / SOM_{s1}^c, m / SOM_{s2}^c, m / SOM_{s3}^c] z_5$

Tabla 3.1 Tabla de decisión normalizada y promediada según los criterios y sus pesos asociados

- 7) El séptimo paso consiste en el cálculo de las distancias de cada alternativa a las soluciones ideales positiva y negativa. Esto se realiza por medio de la ecuación (3.52).
- 8) El octavo paso se realiza calculando el coeficiente de proximidad para cada alternativa a través de la ecuación (3.54).
- 9) Finalmente, el noveno y último paso consiste en elegir la mejor alternativa, es decir la clase  $w_j$  más apropiada para la muestra  $x_s$  de entrada a clasificar, que en nuestro caso serán los valores de las componentes espectrales Rojo (R), Verde (G) y Azul (B) en el modelo de color RGB de los píxeles en las imágenes forestales utilizadas. Dado que en nuestro modelo, alternativas y clases son equivalentes, la mejor clase  $w_j$  para  $x_s$  se corresponde con la alternativa que proporciona el máximo valor del coeficiente de proximidad, es decir la elección se hace de acuerdo con la regla siguiente,

$$w_j = \arg \max_j \{ cp_j \} \quad (3.55)$$

# Capítulo 4

## Análisis de resultados

### 4.1 Introducción

A lo largo de este capítulo se presentan tanto los procedimientos de configuración de algoritmos y evaluación de resultados, como los propios resultados obtenidos mediante los algoritmos y métodos detallados en el capítulo tres.

Se detallarán los resultados generados por cada algoritmo individual y finalmente los datos proporcionados por el clasificador híbrido implementado, que constituye la aportación fundamental del trabajo de investigación.

### 4.2 Métodos de evaluación de resultados

Con el fin de hallar la mejor configuración posible para los clasificadores individuales utilizados, se ha seguido el método de la *validación cruzada*, explicado en la sección 4.2.1.

Tras la ejecución de las pruebas pertinentes con cada clasificador, es necesario poder sintetizar los resultados e interpretarlos para poder compararlos con las salidas de otros clasificadores. Para ello se han utilizado imágenes de referencia o “ground truth”, cuya creación y utilización se detallan en la sección 4.2.2.

Finalmente, con el objetivo de comparar los resultados obtenidos, es necesario definir el protocolo a seguir, el cual se especifica en la sección 4.3.3.

#### 4.2.1 Validación cruzada (Cross-validation)

El procedimiento elegido para estimar los parámetros de los clasificadores es el conocido como *validación cruzada* o “*cross-validation*” en la literatura inglesa especializada (Duda y col., 2001). Para llevarlo a cabo se procede como sigue.

Inicialmente se toman varios conjuntos de datos, esto es píxeles diferentes, en nuestro caso se eligieron cinco conjuntos, algunos de los cuales se tomaron con la imagen original, y otros con la imagen habiendo aplicado el filtrado homomórfico. Estos conjuntos los denominamos Muestras de Aprendizaje o  $MA_i$ .

Cada uno de los conjuntos que describen las cinco categorías de datos se utiliza como conjunto de *entrenamiento*. Posteriormente, la imagen completa se utilizará como conjunto de *validación*. Se ha hecho así y no con un subconjunto de los datos aleatorios, con el fin de obtener resultados con respecto a todas las texturas de la imagen, lo cual parece arrojar una información más fiable a pesar del tiempo de procesado que requiere.

Para estimar el mejor valor de un parámetro determinado de acuerdo con cada método de clasificación, se varía el valor de dicho parámetro en un rango de valores establecido como posible o válido para ese parámetro. Después, para cada valor del parámetro, se realiza un proceso de aprendizaje con las muestras pertenecientes al conjunto de entrenamiento, tras lo cual se procede a la clasificación de las muestras pertenecientes al conjunto de validación en la correspondiente fase de decisión según el clasificador en cuestión. Los resultados así obtenidos por el clasificador, se comparan con las clasificaciones conocidas del conjunto de validación, gracias a lo cual se puede determinar el error cometido por cada clasificador en cada conjunto de validación. Se obtiene un valor de error promedio con las cinco categorías de datos. Aquél valor del parámetro que produzca el menor porcentaje de error medio es el finalmente elegido.

En las secciones 4.4 y 4.5, se muestran las tablas con los valores calculados mediante este método. La validación cruzada se ha utilizado para obtener las variables de los clasificadores, los límites de los incrementos aleatorios para generar los números triangulares *fuzzy* y los pesos específicos de los clasificadores utilizados en el MCDMF, basándonos inicialmente en los resultados obtenidos para dichos clasificadores individuales.

También se utilizó para la suma de las componentes espectrales de los píxeles que debía considerarse como límite para el color negro a la hora de excluir los bordes de las imágenes, aunque como veremos al final de la sección 4.3.3, este método fue sustituido por el cálculo de la distancia al centro. Este método tiene también un pequeño margen de

error en píxeles, como también comentaremos en la sección 4.3.3, donde también aplicamos validación cruzada.

### 4.2.2 Imágenes de referencia o “Ground truth”

Con el fin de verificar el comportamiento de los diferentes métodos, tanto individuales como combinados, es necesario construir la base de imágenes de referencia o “ground truth” con las que se han de comparar los resultados.

El procedimiento para elaborar cada imagen de referencia en estas imágenes forestales ha de hacerse a mano, ya que incluso con aplicaciones avanzadas de edición fotográfica, es muy complicado automatizar el proceso de generación de tales imágenes.

No obstante, para la elaboración de estas imágenes es necesario trabajar con un editor que permita la utilización de capas y la posibilidad de hacer ampliaciones. Las imágenes de referencia se generan de forma que las texturas de interés que deben ser evaluadas se etiquetan con el valor lógico de cero, esto es en negro, mientras que las restantes con el valor lógico de uno, o sea blanco. Esto va a permitir la evaluación automática de los resultados por el procedimiento descrito más adelante. En cualquier caso, en nuestros experimentos los píxeles etiquetados en negro son los troncos por ser las texturas de interés.

A continuación, en las figuras 4.1 y 4.2 se muestran imágenes forestales y sus correspondientes imágenes de referencia para un bosque de pinos y eucaliptos respectivamente:

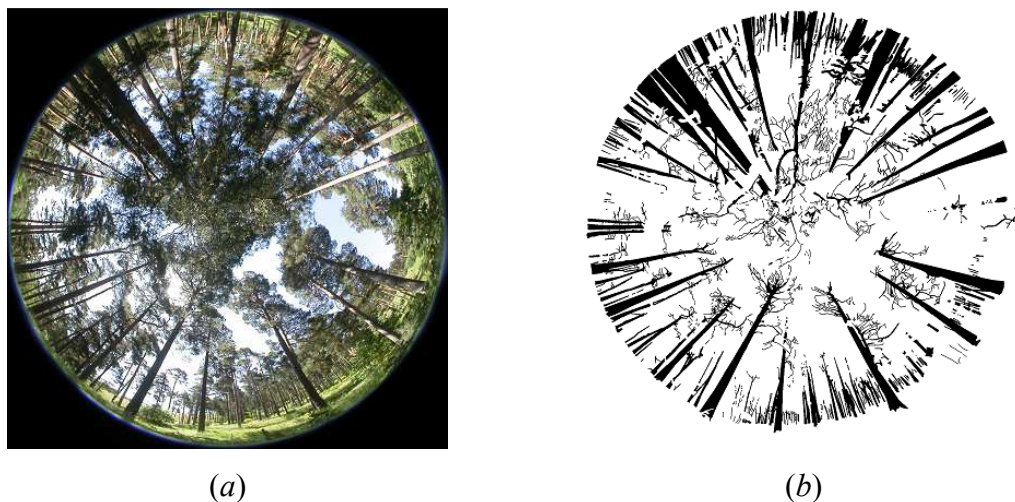


Figura 4.1 (a) Imagen original de bosque de pinos; (b) ground truth correspondiente

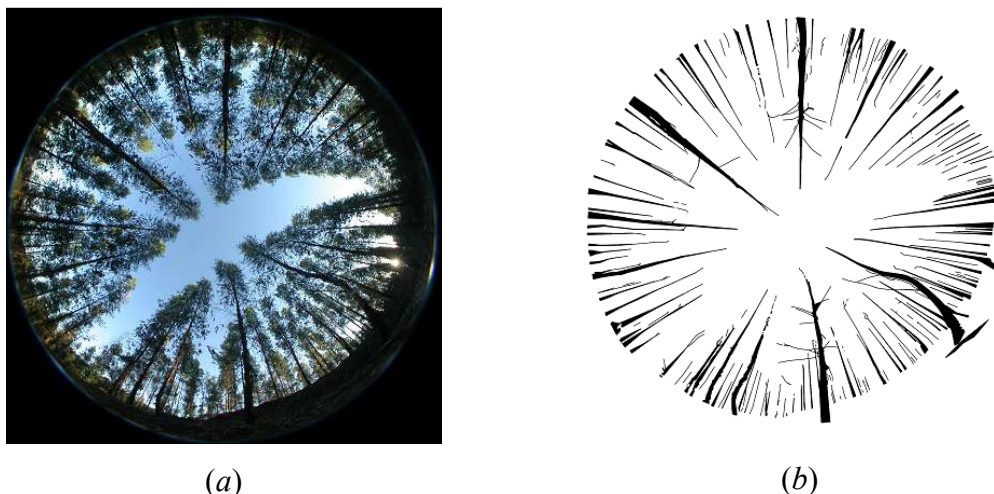


Figura 4.2 (a) Imagen de bosque de eucaliptos; (b) ground truth correspondiente

Cada clasificador, incluido el híbrido, proporciona la correspondiente imagen de salida, donde los píxeles clasificados se etiquetan en negro y los no clasificados en blanco. De esta forma es posible comparar con las imágenes de referencia que poseen esta misma nomenclatura.

Centrándonos sobre el clasificador híbrido, una vez se obtiene la salida del mismo, se recorre la imagen y se crea una máscara píxel a píxel con la imagen de ground truth, anotando tres valores diferentes:

- El número total de píxeles negros en la imagen de referencia, al que denominaremos  $N$ .
- El número total de coincidencias de la clase o clases que el clasificador híbrido define como troncos con píxeles negros en la imagen de referencia. A este valor lo llamaremos  $A$ .
- El número total de errores, esto es, que habiendo sido identificados por el clasificador híbrido como troncos, no se corresponden con un píxel negro en la imagen de referencia, a partir de ahora  $E$ .

Pues bien, de estos valores obtenemos dos porcentajes diferentes, que son:

- *Porcentaje total de troncos*: Porcentaje total de aciertos o coincidencias de la clase o clases de los troncos con respecto al número ideal de píxeles en la imagen de referencia, es decir, qué cobertura se hace exactamente a los troncos.

$$t = 100 \cdot A / N \quad (5.1)$$



- *Porcentaje total de aciertos*: La diferencia con respecto al porcentaje anterior es que restamos de los aciertos el número de píxeles considerados como pertenecientes a los troncos que en realidad no lo son.

$$a = 100 \cdot (A - E) / N \quad (5.2)$$

En la figura 4.3 se representa de izquierda a derecha, el resultado obtenido por la salida del clasificador Agrupamiento borroso, las dos clases que en esta imagen han sido consideradas como troncos, y la cobertura que realizan estas clases en la imagen de referencia.

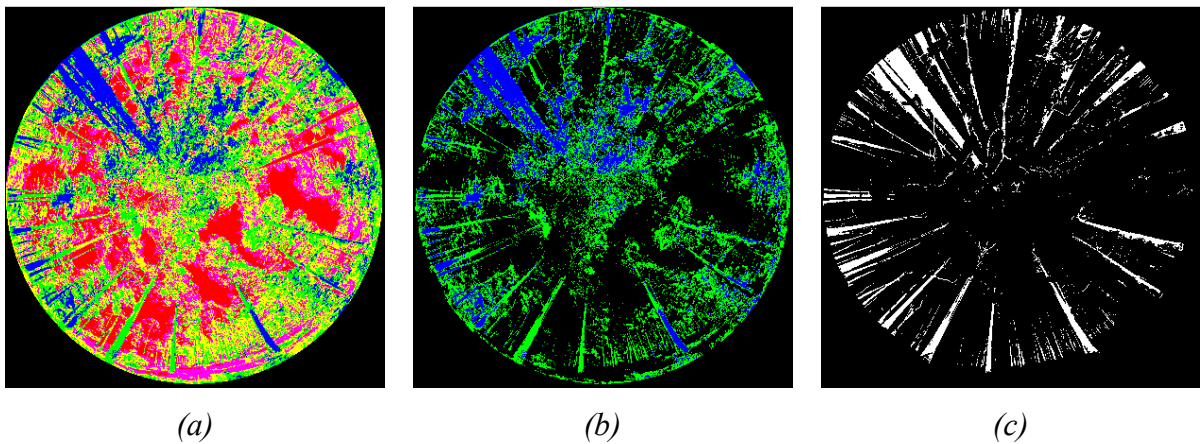


Figura 4.3 Proceso de evaluación del resultado obtenido por un clasificador individual

En el caso concreto de la imagen representada en la figura 4.3, el valor de  $t$  según la ecuación (5.2) resultó ser de 65,5 % y  $a$  según la ecuación (5.2) resultó -86,18%. Podemos observar en esta imagen cómo se pone de manifiesto la confusión entre las texturas de las ramas y el follaje de los pinos, tal como explicaremos con respecto a esta misma imagen en la sección 4.3.2. Este tipo de confusiones es muy habitual en este tipo de imágenes, lo cual se traduce en un empeoramiento importante del porcentaje de aciertos a pesar de que la cobertura de troncos sobre el ground truth no sea aparentemente mala.

Es importante tener en cuenta que las imágenes de referencia están obtenidas a mano por un operador humano sujeto a los errores propios, y por tanto no son precisas al cien por cien, puesto que en ocasiones es complicado diferenciar las texturas incluso para el ojo humano. Por tanto los errores en la evaluación deben considerarse con un cierto grado de flexibilidad.

En la sección siguiente se especifica el criterio seguido para determinar la calidad de los porcentajes, es decir, qué valores se han considerado mejores, teniendo en cuenta ambos simultáneamente.

### **4.2.3 Criterio de calidad de los porcentajes**

A la hora de evaluar los porcentajes obtenidos por los diferentes algoritmos, debemos recordar cuál es el objetivo de este trabajo. La intención es eliminar el máximo posible de texturas no identificadas como troncos, de modo que se facilite el tratamiento posterior de las imágenes para su procesamiento mediante técnicas de visión estereoscópica.

Es por esto por lo que es esencial que se cubra el máximo posible de los troncos, de modo que la imagen resultado mantenga la presencia de éstos para su tratamiento en estereoscopía.

Por otro lado, es importante que el algoritmo cometa el menor número de errores posible, ya que si se cubren todos los troncos pero se suponen como troncos otras muchas texturas, estaríamos entorpeciendo la labor posterior.

Con lo cual, debemos encontrar un compromiso entre el porcentaje de cobertura de los troncos, definido como  $t$ , y el porcentaje de aciertos  $a$ , definidos ambos en la sección 4.2.2.

Llegados a este punto, sería deseable que se consiguiera al menos un 80 % de cobertura de los troncos, con el mayor porcentaje de aciertos posible, no obstante, como los troncos realmente interesantes para el procesado posterior son los más próximos al sistema de captura de imágenes, esto es la cámara, se podrían también llegar a considerar buenos porcentajes de cobertura valores entre el 60% y el 70%, ya que muchos troncos en el fondo de las imágenes no serían evaluados durante el posterior proceso de visión estereoscópica. De hecho, únicamente se evalúan aquellos troncos que se encuentran entre los 5 y los 25m de distancia con respecto al dispositivo de captura, siendo esta una restricción impuesta por lo técnicos forestales, debido a que los árboles situados más allá de esta distancia no se puede garantizar la exactitud de los resultados obtenidos.

Los resultados obtenidos se detallan en la sección 4.4 para los clasificadores individuales y en la sección 4.5 para el clasificador híbrido.

Adelantando,  $a$  siempre tomará un valor mucho menor que  $t$ , puesto que es prácticamente imposible que el algoritmo no cometa errores en este tipo de imágenes. La medida de calidad utilizada es la definida a continuación:

Se considera que un par de porcentajes  $(a, t)$  es mejor que otro par  $(a', t')$  si se cumple, que, siendo  $t$  y  $t'$  mayores del 60%,  $t - a < t' - a'$ .

Si  $t$  es menor del 60%, podemos afirmar que el algoritmo no cubre los troncos suficientes y necesarios para poder realizar un procesamiento posterior satisfactorio de la imagen, con lo que un par  $(a, t)$  será peor cuanto menor sea  $t$ .

### 4.3 Características de las imágenes tratadas

Una vez descritos los procedimientos que seguiremos para obtener los parámetros de los algoritmos y evaluar los resultados, realizamos un breve estudio de las imágenes, con el fin de seleccionar una imagen con características intermedias de entre la base de datos disponible.

En esta sección evaluaremos las características de las imágenes tales como la iluminación o la heterogeneidad de las texturas, para finalizar explicando los métodos de preprocesado de las imágenes realizados con el fin de simplificar la tarea de los clasificadores.

#### 4.3.1 Iluminación

Tal como comentábamos en el capítulo uno, estas imágenes están tomadas en entornos naturales y por tanto de exterior no estructurados, con lo cual las condiciones del entorno resultan determinantes en los resultados.

Una de las características que más influye a la hora de que un algoritmo de tratamiento de imágenes funcione correctamente es la iluminación. En la figura 4.4 vemos cómo en (a), el sol distorsiona un grupo de troncos haciendo prácticamente imposible su reconocimiento. En (b) sin embargo, cómo sin llegar a distorsionar la imagen, existen troncos demasiado iluminados que pueden llegar a confundirse con la textura que forma el cielo. Finalmente, en (c) vemos una imagen tomada al anochecer, en la que es prácticamente imposible diferenciar los troncos en su entorno natural.

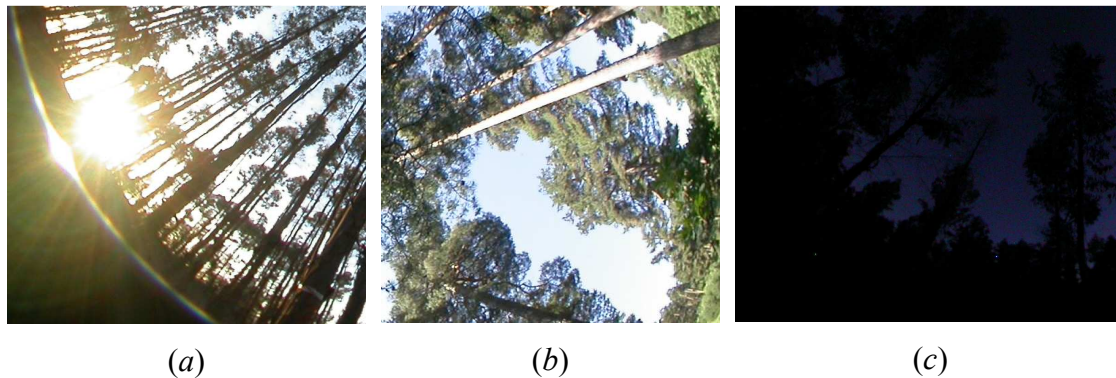


Figura 4.4 Diferencias en la iluminación en diferentes imágenes forestales

Estas características pueden hacer y de hecho hacen que el algoritmo de tratamiento de imágenes basado únicamente en las características espectrales de los píxeles de las imágenes, no sea suficiente para determinar dónde se encuentran los troncos en las imágenes con exactitud para así poder identificarlos y ubicarlos en la propia imagen.

Para tratar de paliar los problemas con la iluminación, se han utilizado dos técnicas conocidas de procesamiento de imágenes, que se aplican en la fase de segmentación. Estas técnicas se detallan en la sección 4.3.3, y son:

- *Realzado*: para imágenes oscuras y poco iluminadas.
- *Filtrado homomórfico*: para imágenes con iluminación variable, es decir, con partes de la imagen iluminadas y otras ensombrecidas.

### 4.3.2 Complejidad de las texturas

Otra de las características que dificulta la extracción de las texturas de los troncos en este caso, es la heterogeneidad con la que se caracterizan este tipo de texturas naturales. En ciertos casos esta variabilidad en una misma textura se debe a cambios en la iluminación, tal como comentábamos en la sección anterior, y podemos observar de nuevo en la figura 4.5. En esta imagen de un bosque de pinos, se observa cómo la iluminación provoca que el tronco más relevante sea prácticamente indistinguible del cielo al menos en la parte inferior de éste. En este caso lo que ha ocurrido es un fenómeno de saturación del sensor CCD de la cámara. Esto es, resulta tan potente la iluminación en dicha zona, que provoca en los píxeles afectados un brillo tal que hace que el sensor en esos píxeles alcance su máximo valor, en este caso de 255 en la representación de 8 bits como es el caso. Esta

saturación se produce en las tres componentes espectrales RGB haciendo que esos píxeles alcancen el blanco puro, es decir (255, 255, 255).



Figura 4.5 Tronco iluminado que presenta texturas similares a las de cielo

En la figura 4.6 sin embargo, se produce el efecto contrario al anteriormente descrito, en efecto en ella observamos que los troncos se confunden con los elementos que los rodean. En (a) los troncos se confunden con el suelo y con el follaje, en (b), el tronco está cubierto por musgo, con lo cual las zonas verdes se confunden con la vegetación del suelo y la madera con el follaje lejano. En (c) se aprecia que los troncos están prácticamente ocultos, y tan sólo podemos intuir vagamente dónde se encuentran, si bien en este último caso por intuición humana, con lo cual las técnicas de clasificación, que carecen de esta capacidad tendrán una dificultad prácticamente insalvable a la hora de establecer las clasificaciones respecto a tales troncos.

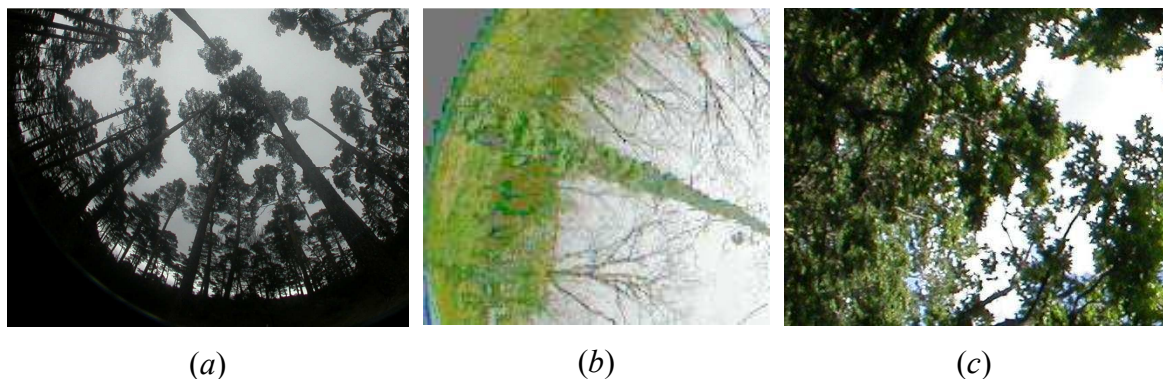


Figura 4.6 Troncos con texturas complicadas

En ocasiones, la imagen posee más elementos con la misma textura que los troncos, que no interesarían para un procesamiento posterior. En la figura 4.7, las ramas existentes en la parte central están entremezcladas con las hojas de los pinos, pequeñas pero frondosas,

la consecuencia de esto es el elevado número de fallos que se genera en este caso, haciendo que el porcentaje de aciertos “ $a$ ” disminuya considerablemente.



Figura 4.7 Múltiples ramas con la misma textura que los troncos

### 4.3.3 Preprocesado de las imágenes

En esta sección se especifican las técnicas de preprocesado introducidas en la sección 4.3.1. Como comentábamos anteriormente, para imágenes con poca iluminación utilizaremos la técnica del realzado y para imágenes con iluminación irregular utilizaremos la técnica de filtrado homomórfico.

Adicionalmente y para finalizar con esta sección, explicaremos la sencilla técnica utilizada para excluir el borde o marco de la imagen.

- *Realzado:*

El realzado aplicado consiste simplemente en la técnica de igualación del histograma sobre la componente de intensidad  $I$ , obtenida por la transformación de la imagen desde el modelo de color RGB al HSI (Pajares y Cruz, 2007). En la figura 4.8(a) se muestra una imagen original procedente de un bosque de robles obtenida con una iluminación relativamente escasa. Con esta sencilla técnica y realizando la transformación inversa, una vez realzada la imagen de intensidad, desde el modelo HSI al RGB se obtiene el resultado mostrado en la imagen de la figura 4.8(b).



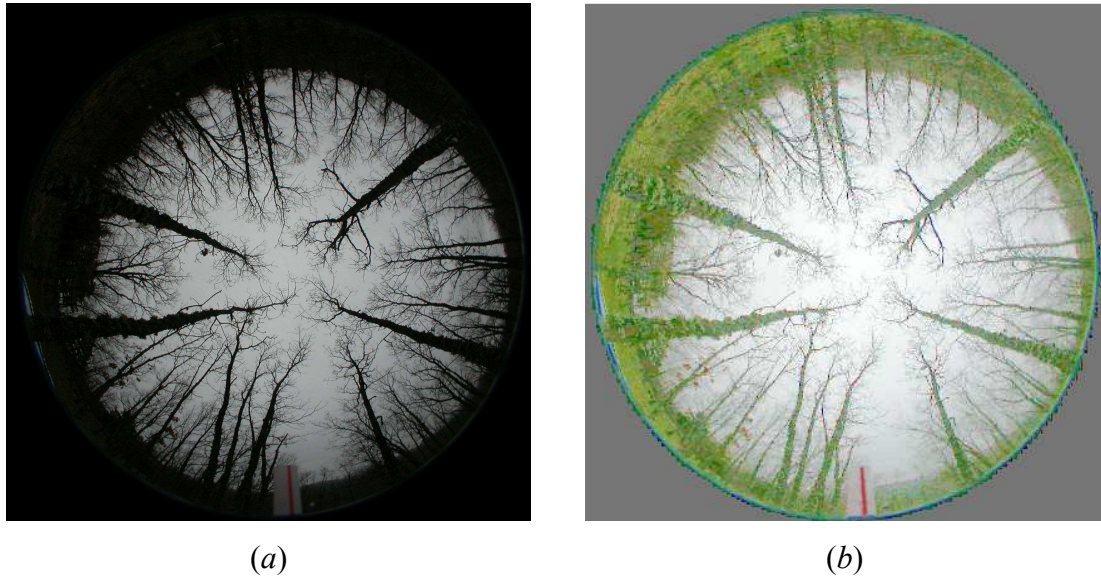


Figura 4.8 (a) Imagen original con escasa iluminación; (b) realzado mediante igualación de histograma de la imagen (a)

En la figura 4.9 comprobamos cómo seleccionando aproximadamente las mismas muestras iniciales, los píxeles se agrupan en clases con más dispersión en la imagen realzada que en la original. En el caso de la imagen en cuestión esto resulta positivo, ya que se podrán identificar texturas nuevas con más facilidad.

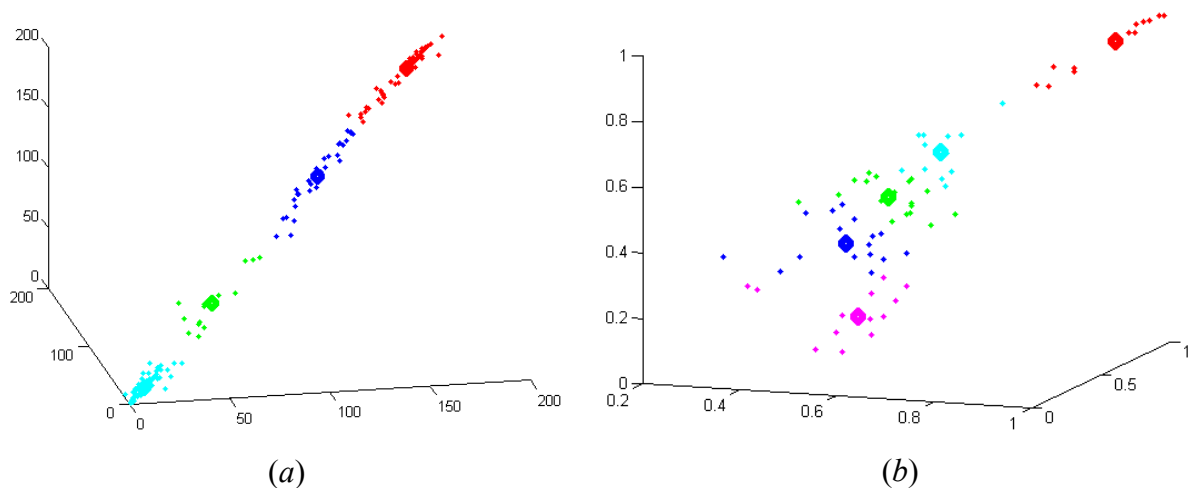


Figura 4.9 (a) Agrupación en clases de píxeles de muestra obtenidos de la figura 4.8(a); (b) agrupación de los mismos píxeles en clases a partir de la figura 4.8(b)

- *Filtrado homomórfico:*

El filtrado homomórfico funciona para imágenes con ruido producido por la variación de iluminación (Pajares y Cruz, 2007).

Como estamos tratando con imágenes tomadas en pares estereoscópicos, tomamos una de ellas y conociendo las características de frecuencia de ésta, aproximamos la segunda imagen a las características espectrales HSI, substraemos la iluminación, relativizándola a la de la primera imagen.

De este modo obtenemos una imagen en la que se ha homogenizado la iluminación, eliminando así la variabilidad, con lo que se reducen los efectos negativos de esta variación en el procesado posterior a la clasificación.

En la figura 4.10 podemos observar los efectos del filtrado homomórfico en una imagen de un bosque de pinos. Esta imagen es la misma que la utilizada en la figura 4.7.

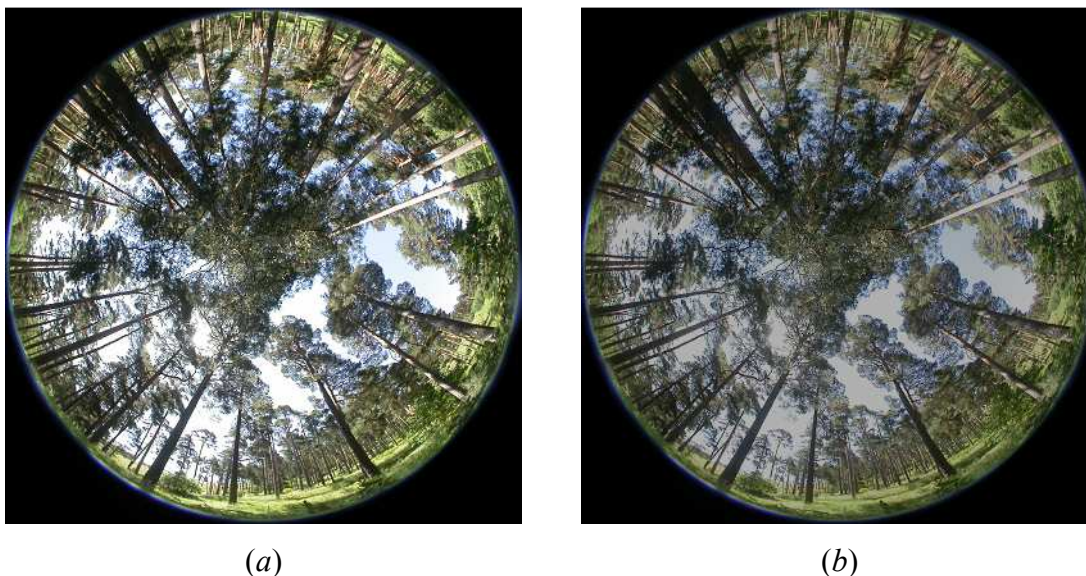


Figura 4.10 (a) Imagen original con iluminación irregular; (b) filtrado homomórfico sobre (a)

En este caso el impacto con respecto a la clasificación inicial es mínimo. La mejora de los resultados se hace evidente en la clasificación de nuevos patrones, ya que éstos son más homogéneos.

- *Exclusión de bordes:*

Se trata de excluir píxeles que no necesitan ser procesados por los clasificadores. En nuestro caso resulta evidente que el borde no debe procesarse.

En un principio se optó por excluir los píxeles de color negro, pero el problema es que ciertos píxeles cercanos al borde no eran 100% negros, sino que presentaban alguna



pequeña componente de color. Se consideraban negros por lo tanto los píxeles cuyas tres componentes de color no sumaran más de una cantidad determinada. Variando este valor se obtenían diferentes resultados, pero cuanto más se incrementaba, más píxeles se excluían de las texturas oscuras, lo cual tampoco es deseable. También se desechó esta técnica ya que al preprocesar la imagen, los píxeles negros se alteraban notablemente, afectando de forma considerable a numerosos píxeles pertenecientes a texturas de interés en las imágenes.

Se optó por cambiar a la técnica que se utiliza actualmente, de forma que para excluir un píxel, se mide su distancia al centro y se compara con el radio de la circunferencia. Este radio está prefijado puesto que todas las imágenes tienen la misma dimensión, aunque se introduce un pequeño margen de variabilidad, que se ha configurado a 5 píxeles, ya que es el valor para el que se han obtenido los mejores resultados en las pruebas de *validación cruzada*. Si la distancia al centro es mayor que el radio, entonces se excluye el píxel que verifica esta condición, con lo cual el proceso se ha simplificado considerablemente al aplicar este sencillo criterio geométrico.

#### 4.3.4 Determinación del número inicial de clases

Como se ha definido en el capítulo tres, el clasificador híbrido calcula inicialmente el número de clases necesario mediante un algoritmo, que tras cada iteración aplica un sencillo cálculo que valida la partición generada hasta el momento. Si las clases no son lo suficientemente buenas, según dicho criterio, se incrementa el número de clases en una unidad y se reinicializa la generación de clases iniciales.

El algoritmo de validación empleado, utiliza una combinación del coeficiente de partición (PC) y la entropía de partición (PE), explicados en la sección 3.3, de modo que si PC es mayor que un valor determinado  $max\_PC$  y PE es menor que  $min\_PE$ , entonces la partición se da como válida.

Para fijar  $max\_PC$  y  $min\_PE$ , primeramente se ha calculado el número óptimo de particiones iniciales de forma manual, con el fin de intentar fijar estos valores para obtener los mejores resultados posibles variando los parámetros de los diferentes algoritmos.

De este modo, como PC y PE están asociados al clasificador Agrupamiento borroso, se han realizado pruebas de validación cruzada forzando el número inicial de clases y el máximo de clases a 3, 4, 5 y 6. Los mejores resultados obtenidos, tal como veremos en la sección 4.4.1, se obtuvieron con 5 clases, con un max\_PC de 0.85 y un min\_PE de 0.15.

Para conseguir validar una partición de más de seis clases, era necesario relajar demasiado el criterio, y a simple vista, no sería deseable que en la imagen se clasificaran más de seis texturas, con lo cual siempre se ha fijado como límite el valor de seis clases.

No obstante, el método de validación cruzada empleado para hallar los parámetros de configuración de los diferentes clasificadores, se ha probado con diferente número de clases por interés en los resultados que podría arrojar.

Llegados a este punto, podemos configurar de nuevo el número de clases inicial en uno y dar libertad al algoritmo, con un máximo de seis clases, como comentábamos anteriormente, para que estime el número de clases óptimo según el criterio establecido a través de los coeficientes de validación de las particiones.

## **4.4 Resultados de los algoritmos individuales**

En esta sección se muestran, los resultados más representativos de uno de los cinco conjuntos de datos iniciales para el que mejor se comportaron los clasificadores en promedio. Dichos conjuntos para calcular los parámetros de los clasificadores con el método de validación cruzada, definidos como muestras de aprendizaje en la sección 4.2.1, se han obtenido de la imagen de pinos que mostrábamos en la figura 4.10, utilizando posteriormente imágenes de pinos similares para obtener los resultados.

Se han elegido las imágenes de los pinos puesto que se ha considerado que reúnen la mayoría de las características expuestas en la sección 4.3 de este mismo capítulo.

En las tablas donde se recogen los resultados podemos observar, tanto los valores de configuración de los algoritmos que se utilizarán posteriormente en el clasificador híbrido, como los porcentajes de acierto y cobertura de troncos descritos en la sección 4.2.2, resaltados con sombreado gris. Los valores de los parámetros de configuración para cada uno de los clasificadores, que se especifican a lo largo del capítulo tres durante la

definición de cada clasificador utilizado, se obtienen mediante el método de validación cruzada detallado en la sección 4.2.1.

La comparación de los resultados obtenidos en esta sección con los del clasificador híbrido, se detalla en la sección 4.5.

#### 4.4.1 Agrupamiento Borroso

La tabla de resultados 4.1 muestra en la primera columna el número de clases; en las columnas dos y tres los parámetros específicos del clasificador de agrupamiento borroso visto en la sección 3.2.1; en las columnas cuarta y quinta, los valores límite para las funciones criterio para la partición, explicados en la sección 3.3; en la sexta columna se detalla el número de pixels para el margen del marco, como comentábamos en la sección 4.3.3; en la séptima columna, si se preprocesó la imagen utilizando la técnica de filtrado homomórfico o no, visto también en la sección 4.3.3, y en las tres últimas columnas, los porcentajes de resultados y la valoración calculada con el método definido en la sección 4.2.3.

Clases	Peso exponencial (m)	Criterio de Terminación	PC	PE	Margen marco	Homfilter	% aciertos (a)	% troncos (t)	Valoración
3	2,1	0,1	0,85	0,15	5	NO	-169,93	73,93	243,86
3	2,3	0,1	0,8	0,2	5	NO	-168,39	73,72	242,11
3	2,5	0,1	0,75	0,25	5	NO	-152,27	71,29	223,56
3	1,9	0,1	0,85	0,15	5	NO	-143,72	73,98	217,7
3	1,9	0,1	0,85	0,15	6	NO	-138,59	73,91	212,5
3	1,9	0,1	0,85	0,15	7	NO	-134,61	73,49	208,1
4	1,9	0,1	0,85	0,15	5	NO	-46,15	50,35	96,5
4	1,9	0,01	0,85	0,15	5	NO	-47,58	50,83	98,41
4	1,9	0,001	0,85	0,15	5	NO	-47,79	50,93	98,72
4	1,9	0,0001	0,85	0,15	5	NO	-47,79	50,97	98,76
4	1,9	0,00001	0,85	0,15	5	NO	-47,80	50,98	98,78
5	1,9	0,00001	0,85	0,15	5	NO	-180,54	78,82	259,36
<b>5</b>	<b>1,9</b>	<b>0,00001</b>	<b>0,85</b>	<b>0,15</b>	<b>5</b>	<b>SI</b>	<b>-86,18</b>	<b>62,5</b>	<b>148,68</b>
3	1,9	0,1	0,85	0,15	5	SI	-113,62	68,45	182,07
4	1,9	0,00001	0,85	0,15	5	SI	-57,42	53,80	111,22
6	1,9	0,00001	0,85	0,19	5	SI	-91,41	63,81	155,22

Tabla 4.1 Valores obtenidos para el clasificador de Agrupamiento borroso

#### 4.4.2 Estimación no paramétrica: ventana de Parzen

La tabla de resultados 4.2 muestra en la primera columna el número de clases; en la segunda columna, si se preprocesó la imagen utilizando la técnica de filtrado homomórfico o no, visto en la sección 4.3.3; en la tercera columna, el parámetro de ancho de banda explicado en la sección 3.2.3, y en las tres últimas columnas, los porcentajes de resultados y la valoración calculada con el método definido en la sección 4.2.3.

Clases	Homfilter	Ancho de banda (h)	% aciertos (a)	% troncos (t)	Valoración
<b>5</b>	<b>SI</b>	<b>4</b>	<b>-66,46</b>	<b>60,13</b>	<b>126,59</b>
5	SI	2	-159,66	78,08	237,74
6	SI	4	-123,32	68,80	192,12
6	SI	2	-165,14	75,41	240,55
4	SI	4	-36,11	48,45	84,56
4	SI	3	-101,63	68,53	170,16
4	SI	2	-143,74	76,09	219,83
3	SI	2	0,53	6,52	5,99

Tabla 4.2 Valores obtenidos para el clasificador de ventana de Parzen

#### 4.4.3 Clasificador Paramétrico: Bayesiano

La tabla de resultados 4.3 muestra en la primera columna el número de clases; en la segunda columna, si se preprocesó la imagen utilizando la técnica de filtrado homomórfico o no, visto en la sección 4.3.3, y en las tres últimas columnas, los porcentajes de resultados y la valoración calculada con el método definido en la sección 4.2.3.

Clases	Homfilter	% aciertos (a)	% troncos (t)	Valoración
<b>5</b>	<b>SI</b>	<b>-75,88</b>	<b>60,93</b>	<b>136,81</b>
6	SI	-183,07	77,99	261,06
4	SI	-55,74	47,68	103,42
3	SI	-87,1	61,79	148,89

Tabla 4.3 Valores obtenidos para el clasificador Bayesiano

#### 4.4.4 Algoritmo Generalizado de Lloyd

La tabla de resultados 4.4 muestra en la primera columna el número de clases; en la segunda columna, si se preprocesó la imagen utilizando la técnica de filtrado

homomórfico o no, visto en la sección 4.3.3; en las columnas tres a cinco, los parámetros para el clasificador basado en el algoritmo generalizado de Lloyd, explicado en la sección 3.2.4, y en las tres últimas columnas, los porcentajes de resultados y la valoración calculada con el método definido en la sección 4.2.3.

Clases	Homfilter	Razón Aprendizaje	Max Iteraciones	Tolerancia	% aciertos (a)	% troncos (t)	Valoración
5	SI	0,1	1000	1,00E-10	-92,94	64,41	157,35
5	SI	0,2	1000	1,00E-10	-155,98	74,87	230,85
5	SI	0,3	1000	1,00E-10	-155,58	75,18	230,76
5	SI	0,65	1000	1,00E-10	-121,34	71,18	192,52
5	SI	0,05	1000	1,00E-10	-84,06	62,11	146,17
5	SI	0,05	5000	1,00E-10	-84,06	62,11	146,17
5	SI	0,05	100	1,00E-10	-84,06	62,11	146,17
5	SI	1	1000	1,00E-10	-116,06	70,87	186,93
5	SI	1	5000	1,00E-20	-116,06	70,87	186,93
6	SI	0,05	1000	1,00E-10	-115,64	69,03	184,67
<b>6</b>	<b>SI</b>	<b>1</b>	<b>1000</b>	<b>1,00E-10</b>	<b>-76,07</b>	<b>61,72</b>	<b>137,79</b>
4	SI	0,05	1000	1,00E-10	-65,22	56,32	121,54
4	SI	1	1000	1,00E-10	-153,09	75,83	228,92
3	SI	0,05	1000	1,00E-10	-60,22	54,73	114,95
3	SI	1	1000	1,00E-10	-60,45	54,29	114,74

Tabla 4.4 Valores obtenidos para el clasificador del algoritmo generalizado de Lloyd

#### 4.4.5 Mapas Auto-Organizativos

La tabla de resultados 4.5 muestra en la primera columna el número de clases; en la segunda columna, si se preprocesó la imagen utilizando la técnica de filtrado homomórfico o no, visto en la sección 4.3.3; en las columnas tres a siete, los parámetros para el clasificador basado en mapas auto-organizativos, explicado en la sección 3.2.5, y en las tres últimas columnas, los porcentajes de resultados y la valoración calculada con el método definido en la sección 4.2.3.

Clases	Homfilter	$\alpha_i$	$\alpha_f$	Nº Max. Iter.	Umbral	Tolerancia	% aciertos (a)	% troncos (t)	Valoración
5	SI	1,00E-01	1,00E-02	1000	1,00E-05	1,00E-06	-65,83	56,32	122,15
5	SI	1,00E-01	1,00E-02	1000	1,00E-05	1,00E-04	-64,68	55,99	120,67
5	SI	1,00E-01	1,00E-02	1000	1,00E-05	1,00E-08	-65,81	56,3	122,11
5	SI	1,00E-02	1,00E-02	1000	1,00E-05	1,00E-06	-59,88	54,55	114,43

5	SI	1,50E-01	1,00E-02	1000	1,00E-05	1,00E-06	-69,41	57,48	126,89
5	SI	1,50E-01	1,00E-03	1000	1,00E-05	1,00E-06	-68,23	57,14	125,37
5	SI	1,50E-01	1,00E-05	1000	1,00E-05	1,00E-06	-67,57	58,96	126,53
<b>5</b>	<b>SI</b>	<b>1,50E-01</b>	<b>1,00E-05</b>	<b>5000</b>	<b>1,00E-05</b>	<b>1,00E-06</b>	<b>-70,7</b>	<b>60,92</b>	<b>131,62</b>
5	SI	1,50E-01	1,00E-05	500	1,00E-05	1,00E-06	-66,28	56,52	122,8
5	SI	1,50E-01	1,00E-05	2000	1,00E-05	1,00E-06	-68,28	57,15	125,43
5	SI	1,50E-01	1,00E-05	1000	1,00E-03	1,00E-06	-67,57	56,95	124,52
6	SI	1,00E-01	1,00E-02	1000	1,00E-05	1,00E-06	-139,92	73,18	213,1
6	SI	1,50E-01	1,00E-05	1000	1,00E-05	1,00E-06	-127,71	71,23	198,94
4	SI	1,00E-01	1,00E-02	1000	1,00E-05	1,00E-06	-181,09	78,76	259,85
4	SI	1,50E-01	1,00E-05	1000	1,00E-05	1,00E-06	-191,19	79,99	271,18
3	SI	1,00E-01	1,00E-02	1000	1,00E-05	1,00E-06	-40,4	47,57	87,97
3	SI	1,50E-01	1,00E-05	1000	1,00E-05	1,00E-06	-33,22	43,79	77,01

Tabla 4.5 Valores obtenidos para el clasificador de mapas auto-organizativos

## 4.5 Resultados del clasificador híbrido

Antes de mostrar los resultados obtenidos por el clasificador híbrido con el fin de compararlos con los valores de las tablas de la sección 4.4, debemos hacer mención a la configuración final, a alto nivel, del clasificador híbrido.

Para realizar las pruebas finales, se prescindió de la funcionalidad de aprendizaje incremental, de la que hablábamos al final de la sección 3.1.1, puesto que en las pruebas iniciales, introducía demasiada variabilidad a las clases y hacía demasiado complicada la validación de estas en ejecuciones sucesivas.

El aprendizaje incremental complica también el coste del algoritmo, puesto que cuando se clasifica cada muestra, es necesario reentrenar los diferentes clasificadores, y debemos tener en cuenta que las imágenes tratadas tienen un volumen de información importante.

Es por esto por lo que se prescinde en las pruebas del módulo de código que efectúa el mencionado aprendizaje incremental.

La tabla de resultados 4.6 muestra en la primera columna el número de clases; en las columnas dos a seis los pesos específicos configurados para cada uno de los clasificadores; en la séptima el valor aleatorio máximo para clasificadores que devuelven porcentajes, ya sea probabilidades o grados de pertenencia; en la octava el valor aleatorio para los clasificadores que devuelven distancias, y finalmente en las columnas nueve a

once los porcentajes de resultados y la valoración calculada con el método definido en la sección 4.2.3.

Como se puede comprobar, se han realizado pruebas con los pesos específicos igualados y eligiéndolos en función de los resultados obtenidos para los clasificadores individuales, de modo que el mejor valorado para cada número de clases en particular obtiene peso 5/15, el siguiente 4/15 y así hasta llegar al peor, cuyo peso es de 1/15.

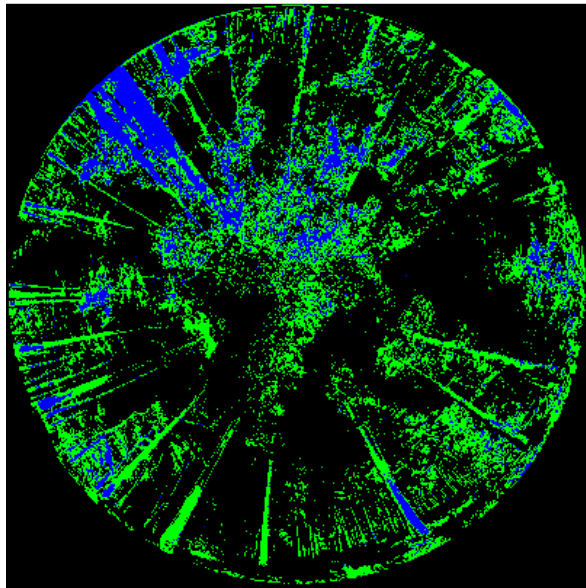
Clases	z1 (FC)	z2 (PB)	z3 (PZ)	z4 (AGL)	z5 (SOM)	random (%)	random (dist)	% aciertos (a)	% troncos (t)	Valoración
5	1/5	1/5	1/5	1/5	1/5	0,1	20	-72,22	61,43	133,65
5	1/15	3/15	5/15	2/15	4/15	0,05	10	-74,73	60,94	135,67
<b>5</b>	<b>1/15</b>	<b>3/15</b>	<b>5/15</b>	<b>2/15</b>	<b>4/15</b>	<b>0,1</b>	<b>20</b>	<b>-63,22</b>	<b>62,61</b>	<b>125,83</b>
5	1/15	3/15	5/15	2/15	4/15	0,2	50	-68,92	60,96	129,88
5	1/15	3/15	5/15	2/15	4/15	0,1	30	-65,71	61,05	126,76
4	1/5	1/5	1/5	1/5	1/5	0,1	20	-124,79	60,43	185,22
4	2/15	1/15	5/15	4/15	3/15	0,1	20	-101,66	68,37	170,03
3	1/5	1/5	1/5	1/5	1/5	0,1	20	-123,76	54,25	178,01
3	4/15	5/15	1/15	3/15	2/15	0,1	20	-85,72	61,85	147,57
6	1/5	1/5	1/5	1/5	1/5	0,1	20	-125,84	66,78	192,62
6	4/15	1/15	3/15	5/15	2/15	0,1	20	-72,94	62,32	135,26

Tabla 4.6 Valores obtenidos para el clasificador híbrido

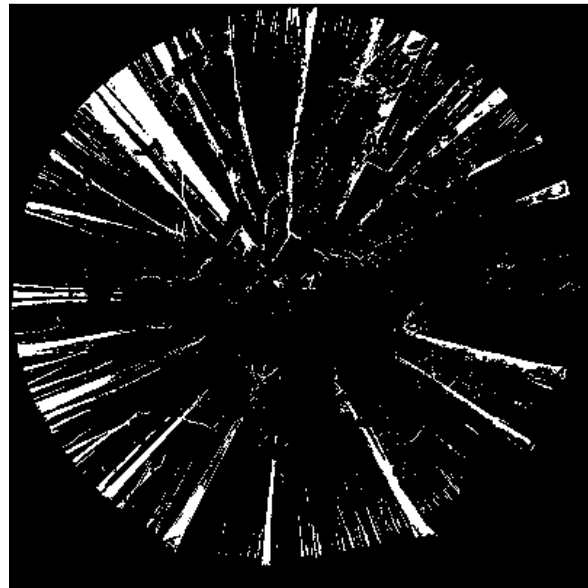
Como se observa a partir de los resultados, los valores obtenidos en el clasificador híbrido son más estables que los obtenidos en los clasificadores individuales. Con los pesos específicos repartidos de forma equitativa, se obtienen valores intermedios, y en algunos casos mejores, que los obtenidos en los clasificadores individuales.

Configurando los pesos específicos y los máximos aleatorios para el algoritmo MCDMF, vemos que los resultados mejoran sensiblemente, consiguiendo en la mayoría de los casos, porcentajes de cobertura de alrededor del 60%, y con valoraciones sensiblemente mejores.

En la figura 4.11 podemos ver de forma cualitativa uno de los mejores resultados obtenidos con las dos imágenes de prueba mostradas, que se corresponde con la fila marcada en verde en la tabla 4.6.



(a)



(b)

Figura 4.11 Resultado obtenido con el clasificador híbrido: (a) Clases de los troncos aisladas; (b) Correspondencia de las clases de los troncos con la imagen de ground truth correspondiente



# Capítulo 5

## Conclusiones y Trabajo futuro

### 5.1 Introducción

Llegados a este punto se hace necesario plantear un balance global sobre las líneas de investigación seguidas, que desembocan en las conclusiones expuestas. Por otra parte, se abren una serie de perspectivas de futuro en relación con los temas de investigación abordados y sus posibilidades de ampliación futuras.

### 5.2 Conclusiones

En esta sección se sintetizan las aportaciones realizadas así como las conclusiones más relevantes que se derivan de ellas. El núcleo central de la investigación ha consistido en abordar la problemática de la clasificación de texturas naturales en imágenes hemisféricas de exterior desde la perspectiva de la combinación de clasificadores, que por otra parte constituye una de las principales líneas de investigación abiertas actualmente en el área, tal y como se pone de manifiesto en la literatura especializada.

Se ha planteado una estrategia local, donde la combinación se lleva a cabo durante la fase de decisión, a partir de las salidas proporcionadas por los cinco clasificadores individuales utilizados en esa misma fase. Dichos clasificadores individuales constituyen la base para la propuesta de combinación que se formula.

Para la estrategia combinada se ha propuesto el procedimiento basado en el paradigma de la Teoría de la Decisión Multicriterio Fuzzy (*MCDMF*), donde el énfasis se pone en la forma de mezclar las salidas de los clasificadores individuales, realizando la combinación píxel a píxel dadas la complejidad y las características de las imágenes analizadas. En este sentido se establecen mecanismos apropiados para determinar las competencias de los clasificadores individuales. El objetivo consiste en otorgar a cada clasificador una cierta importancia relativa a la hora de realizar la combinación de sus salidas.

En relación a los objetivos generales expresados en la sección 1.4 del capítulo uno, cabe destacar que se han cumplido las expectativas previstas. Su grado de cumplimiento se concreta como sigue:

1. Se han estudiado los siguientes métodos de clasificación: Agrupamiento borroso, paramétrico Bayesiano, Ventana de Parzen, Algoritmo generalizado de Lloyd y Mapas Auto-Organizativos.
2. Se han utilizado imágenes hemisféricas correspondientes a diversos tipos de texturas naturales de exterior, por cierto con una elevada complejidad.
3. Se ha propuesto e implementado un método híbrido con una aportación novedosa en el procesado de dichas imágenes, mejorando sensiblemente los resultados obtenidos con los clasificadores clásicos individuales.
4. Se proponen una serie de líneas de investigación futuras.

### 5.3 Trabajo futuro

Aunque el algoritmo híbrido implementado proporciona mejores resultados que los clasificadores clásicos utilizados de forma individual, analizando los resultados y la problemática de las imágenes podemos concluir que es necesario continuar la investigación considerando nuevas alternativas para mejorar la identificación de los troncos y la exclusión de texturas indeseadas en este tipo de imágenes.

Teniendo en cuenta lo anterior y la experiencia acumulada con los conjuntos de imágenes tratadas durante el desarrollo de este trabajo de investigación, a continuación se proponen una serie de consideraciones para establecer las pautas a seguir respecto de las futuras líneas de investigación que se abren en este sentido:

- Estudiar la clasificación representando las imágenes en diferentes modelos de color, tales como *HSI*, *YIQ*, *CMYK* entre otros (Pajares y Cruz, 2007). Esto se plantea con el fin de tratar de conseguir una mayor separación entre los datos pertenecientes a las clases.

- Mejorar los errores de clasificación, desarrollando métodos que incorporen información de las texturas. En esta línea se sitúan por ejemplo técnicas estadísticas para obtener atributos de primer y segundo orden (varianza, momentos de orden tres). También considerar la posibilidad de utilizar otros descriptores de texturas tales como filtros de Gabor, *wavelets* (Pajares y Cruz, 2007).
- Elaboración de una base de datos o “*ground-truth*” más elaborada partiendo de la generación de datos sintéticos y búsquedas en Internet.
- Estudio de los problemas relacionados con la iluminación, ya que las imágenes se capturan en diferentes días y por tanto bajo diferentes condiciones de iluminación. En esta línea se requiere la profundización en técnicas relativas al filtrado homomórfico o mediante correspondencia de histogramas (Pajares y Cruz, 2007).
- El clasificador *MCDMF* está basado en la teoría de la decisión, por lo que en base a sus buenos resultados, cabe pensar en la utilización de algún otro paradigma basado en dicha teoría, probablemente la teoría de la utilidad propuesta en Ríos-Insúa y col. (2002) podría ser una buena candidata.
- Establecer la clasificación mediante estrategias de índole global, como son los Mapas Cognitivos *Fuzzy* o el método de relajación Probabilista e incluso técnicas de optimización tales como las redes neuronales de Hopfield o la técnica de Enfriamiento Simulado (Pajares y Cruz, 2007).



# Bibliografia

- Aiyer, A., Pyun, K., Huang, Y., O'Brien, D.B., Gray, R.M. (2005). Lloyd Clustering of Gauss mixture models for image compression and classification, *Signal Processing: Image Communication*, 20, 459-485.
- Arivazhagan, S., Ganesan, L. (2003). Texture classification using wavelet transform, *Pattern Recognition Letters*, 24, 1513-1521.
- Atkinson, P.M., Tatnall, A.R.L. (1997). Neural Networks in Remote Sensing, *Int. Journal of Remote Sensing*, 18, 699-709.
- AVIRIS: Airborne Visible InfraRed Imaging Spectrometer (2009)  
<http://aviris.jpl.nasa.gov/html/aviris.overview.html>
- Balasko, B., Abonyi, J., Feil, B. (2006). Fuzzy Clustering and Data Analysis Toolbox for Use with Matlab, Veszprem University, Hungary  
<http://www.fmt.vein.hu/softcomp/fclusttoolbox/FuzzyClusteringToolbox.pdf>
- Barnard, S., Fishler, M. (1982). Computational Stereo, *ACM Computing Surveys*, 14, 553-572.
- Benediktsson, J.A., Kanellopoulos, I. (1999). Classification of multisource and hyperspectral data based on decision fusion. *IEEE Trans. Geoscience Remote Sensing*, 37(3), 1367-1377.
- Benediktsson, J.A., Swain, P.H., Ersoy, O.K. (1990). Neural network approaches versus statistical methods in classification of multisource remote sensing data, *IEEE Trans. Geosci. Remote Sens.*, 28(4), 540-552.
- Bezdek, J.C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Kluwer, Plenum Press, New York.
- Briem, G.J., Benediktsson, J.A., Sveinsson, J.R. (2002). Multiple Classifiers applied to multisource Remote Sensing Data. *IEEE Trans. Geosci. Remote Sens.*, 40(10), 2291-2299.

- Brodatz, P. (1966). Textures: A photographic album for artists and designers, Dover Publications, New York.
- Broek, E.L., Rikxoort, E.M. (2004). Evaluation of color representation for texture analysis, in: Proceedings of the 16th Belgium-Netherlands Artificial Intelligence Conference, R. Verbrugge, N. Taatgen, and L. R. B. Schomaker, (Eds.), Groningen - The Netherlands, 35–42.
- Bruzzone, L., Fernández-Prieto, D., Serpico, S. B. (1999). A neural-statistical approach to multitemporal and multisource remote-sensing image classification. *IEEE Trans. Geosci. Remote Sens.*, 37(11), 1350–1359.
- Campadelli, P., Medici, D., Schettini, R. (1997). Color image segmentation using Hopfield Networks, *Image and Vision Computing*, 15, 161-166.
- Camps-Valls, G., Bruzzone, L. (2005). Kernel-based methods for hyperspectral image classification, *IEEE Trans. Geosci. Remote Sens.*, 43(6), 1351–1362.
- Camps-Valls, G., Gómez-Chova, L., Calpe, J., Soria, E., Martín, J. D., Alonso, L., Moreno, J. (2004). Robust support vector method for hyperspectral data classification and knowledge discovery, *IEEE Trans. Geoscience. Remote Sensing*, 42(7), 1530–1542.
- Canty, M. J. (2006). *Image Analysis, Classification and Change Detection in Remote Sensing: With Algorithms for ENVI/IDL*, CRC/Taylor & Francis.
- Cao, J., Shridhar, M., Ahmadi, M. (1995). Fusion of Classifiers with Fuzzy Integrals. In: *Proc. 3th Int. conf. Document Analysis and Recognition (ICDAR'95)*, 1, (108-111)
- Chamorro-Martínez, J., Galán-Perales, E., Prados-Suárez B., Soto-Hidalgo, J.M. (2007). Perceptually-Based Functions for Coarseness Textural Feature Representation, in: J. Martí, J.M. Benedí. A.M. Mendoça and J. Serrat (Eds.), *Pattern Recognition and Image Analysis*, Springer-Verlag, Berlin, LNCS 4477, Part I, 578-586.
- Chan, J.C.W., Laporte, N., Defries, R.S. (2003). Texture Classification of logged forest in tropical Africa using machine-learning algorithms, *Int. J. Remote Sensing*, 24(6), 1401-1407.

- 
- Chang, M.M., Sezan, M.I., Tekalp, A.M. (1994). Adaptive Bayesian segmentation of color images, *Journal Electronic Imaging*, 3, 404-414.
- Chang, T., Kuo, C.C.J. (1993). Texture analysis and classification with tree structured wavelet transform, *IEEE Trans. Image Processing*, 2(10), 429-441.
- Chanussot, J., Benediktsson, J.A., Fauvel, M. (2006). Classification of Remote Sensing Images from Urban Areas using a Fuzzy Possibilistic Model, *IEEE Geoscience and Remote Sensing Letters*, 3(1), 40-1407.
- Chen, C.T. (2000). Extensions of the TOPSIS for group decision-making under fuzzy environment. *Fuzzy Sets and Systems* 114, 1-9.
- Chen, J., Pappas, T.N., Mojsilovic, A., Rogowitz, B.E. (2005). Adaptive Perceptual Color-Texture Image Segmentation, *IEEE Trans. Image Processing*, 14(10), 1524-1536.
- Cheng, H.D., Jiang, X.H., Sun, Y., Wang, J. (2001). Color image segmentation: advances and prospects, *Pattern Recognition*, 34 (12), 2259–2281.
- Clark, A.A., Thomas, B.T., Campbell, N.W., Greenway, P. (1999). Texture deconvolution for the Fourier-based analysis of non-rectangular regions. *BMVC*, 193-202.
- Cochran, S.D., Medioni, G. (1992). 3-D Surface Description from binocular stereo, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10), 981-994.
- Comaniciu, D., Meer, P. (2002). Mean shift: a robust approach toward feature space analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24, 603–619.
- Comaniciu, D., Meer, P. (1997). Robust analysis of feature spaces: color image segmentation, in: *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, San Juan, PR, 750-755.
- Debeir, O., van den Steen, I., Latinne, P., van Ham, P., Wolff, E. (2002). Textural and contextual land-cover classification using single and multiple classifier systems, *Photogramm. Eng. Remote Sens.*, 68(6), 597–605.
- DeSieno, D. (1988). Adding a conscience to competitive learning. *Proc. IEEE Int. Conf. Neural Networks*, 1, pp. 117-124, San Diego, CA.
-

- Dimap (2009). Digital Image Processing, S.L. <http://www.dimap.es/>
- Du, Y., Chang, C., Thouin, P.D. (2004). Unsupervised approach to color video thresholding, *Optical Engineering*, 43(2), 282-289.
- Duda, R.O., Hart, P.E., Stork, D.S. (2001). *Pattern Classification*. New York: Wiley & Sons.
- Duin, R.P.W. (1976). On the choice of smoothing parameters for Parzen estimators of probability density functions. *IEEE Trans. Computers*, C-25(11), 1175-1179.
- Dunn, D., Higgins, W.E. (1995). Optimal Gabor filters for texture segmentation, *IEEE Trans. Image Processing*, 4(7), 947-964.
- E-Cognition-Definiens (2009) <http://www.definiens.com/>
- ENVI (2009) <http://www.itvis.com/ProductServices/ENVI.aspx>
- ERDAS Imagine (2009) <http://www.erdas.com>
- Escalera, A. (2001). *Visión por Computador: Fundamentos y Métodos*, Prentice-Hall, Madrid.
- Escudero, L.F. (1977). *Reconocimiento de patrones*. Paraninfo, Madrid.
- Fauvel, M., Chanussot, J., Benediktsson, J.A. (2006). Decision Fusion for the Classification of Urban Remote Sensing Images. *IEEE Trans. Geoscience and Remote Sensing*, 44(10), 2828- 2838.
- Ferreiro-Armán, M., Bandeira, L.P.C., Martín-Herrero, J., Pina, P. (2007). Classifiers for vegetation and forest mapping with low resolution multispectral imagery, in: J. Martí, J.M. Benedí, A.M. Mendoça and J. Serrat (Eds.), *Pattern Recognition and Image Analysis*, Springer-Verlag, Berlin, LNCS 4477, Part I, 177-184.
- Frate, F., Del, Pacifici, F., Schiavon, G., Solimini, C. (2007). Use of Neural Networks for Automatic Classification from High-Resolution Images, *IEEE Trans. Geoscience and Remote Sensing*, 45(4), 800- 809.
- Frizzelle, B.G., Moody, A. (2001). Mapping continuous distributions of land cover: A comparison of maximum-likelihood estimation and artificial neural networks, *Photogramm. Eng. Remote Sens.*, 67(6), 693–705.



- Fukunaga, K., Hummels, M.D. (1987). Bayes error estimation using Parzen and K-NN procedures. *IEEE Trans Pattern Anal. Machine Intelligence*, 9(5), 634-643.
- Geman, S., Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Trans. Pattern Analysis Machine Intelligence*, 6, 721-741.
- Giacinto, G., Roli, F., Bruzzone, L. (2000). Combination of neural and statistical algorithms for supervised classification of remote-sensing image, *Pattern Recognition Letters*, 21(5), 385-397.
- Gonzalez, R., Woods, R. (2007). *Digital Image Processing*, Prentice-Hall.
- Gu, X., Zhu, Q. (2006). Fuzzy multi-attribute decision-making method based on eigenvector of fuzzy attribute evaluation space. *Decision Support Systems*, 41, 400-410.
- Guijarro, M. (2007). Diseño de un nuevo clasificador híbrido no supervisado para texturas naturales en imágenes, Proyecto Fin de Master, Departamento de Ingeniería del Software e Inteligencia Artificial, Facultad de Informática, Universidad Complutense de Madrid.
- Guijarro, M. (2009). Combinación de clasificadores para identificación de texturas en imágenes naturales: nuevas estrategias locales y globales, Tesis doctoral, Facultad de Informática, Universidad Complutense de Madrid.
- Haapanen, R., Ek, A.R., Bauer, M.E., Findley, A.O. (2004). Delineation of forest/nonforest land use classes using nearest neighbor methods, *Remote Sensing Environment*, 89(3), 265-271.
- Hanmandlu, M., Madasu, V.K., Vasikarla, S. (2004). A Fuzzy Approach to Texture Segmentation, in: *Proc. of the IEEE International Conference on Information Technology: Coding and Computing (ITCC'04)*, The Orleans, Las Vegas, Nevada, USA, 636-642.
- Hansen, M, Dubayah, R., Defries, R. (1996). Classification trees: an alternative to traditional land cover classifiers, *Int. J. Remote Sensing*, 17(5), 1075-1081.
- Hansen, L.K., Salamon, P. (1990). Neural Network ensembles. *IEEE Trans. Pattern Analysis and Applications*, 12(10), 993-1001.

- Hashem, S., Schmeiser, B. (1995). Improving Model Accuracy Using Optimal Linear Combinations of Trained Neural Networks, *IEEE Transactions on Neural Networks*, 6(3), 792-794.
- Haykin, S. (1994). *Neural Networks: a comprehensive foundation*. Macmillan College Publishing Co., New York.
- Herrera, P.J., (2010). Correspondencia estereoscópica en imágenes obtenidas con proyección omnidireccional para entornos forestales. Tesis doctoral. Facultad de Informática, Universidad Complutense de Madrid.
- Ho, T. K. (2002). Multiple classifier combination: Lessons and the next steps. In A. Kandel and H. Bunke (Eds.), *Hybrid Methods in Pattern Recognition*. World Scientific Publishing, 171-198.
- Hu, R., Damber, R.I. (2008). A “Non Panacea Theorem” for classifier combination. *Pattern Recognition*, 41(8), 2665-2673.
- Huang, C.L. (1999). Pattern image segmentation using modified Hopfield model, *Pattern Recognition Letters*, 13, 345-353.
- Huang, H.J., Hsu, C.N. (2002). Bayesian classification for data from the same unknown class. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 2002, 32(2), 137-145.
- Hwang, C.L., Yoon, K. (1981). *Multiple Attributes Decision Making Methods and Applications*, Springer, Berlin Heidelberg.
- Intergraph (2009) [www.intergraph.com](http://www.intergraph.com)
- ISCAR. (2006). Grupo Ingeniería de Sistemas, Control, Automatización y Robótica. [www.dacya.ucm.es/area-isa](http://www.dacya.ucm.es/area-isa)
- Jain, A.K., Ramaswami, M.D. (1988). Classifier design with Parzen windows. In *Pattern Recognition and Artificial Intelligence toward an integration*, E.S. Geselma and L.N. Kanal (Eds.) Amsterdam, The Netherlands, 211-218.
- Jiji, G.W., Ganesan, L. (2008). A new approach for unsupervised segmentation, *Applied Soft Computing Journal*, doi:10.1016/j.asoc.2008.06.002.
- Jimenez, L., Morales-Morell, A., Creus, A. (1999). Classification of hyperdimensional data based on feature and decision fusion approaches using projection pursuit,

- majority voting, and neural networks. *IEEE Trans. Geosci. Remote Sens.*, 37(3), 1360–1366.
- Kanellopoulos, I., Wilkinson, G.G., Megier, J. (1993). Integration of neural network and statistical image classifier for land cover mapping. *Proc. Geoscience and Remote Sensing Symposium (IGARSS'93). Better Understanding of Earth Environment*, 2(18-21), 511-513.
- Kaufmann, A., Gupta, M.M. (1985). *Introduction to Fuzzy Arithmetic: Theory and Applications*, Van Nostrand Reinhold, New York.
- Kim, D.W., Lee, K.H., Lee, D (2003). Fuzzy Cluster validation index based on inter-cluster proximity. *Pattern Recognition Letters*, 24 (2003) 2561-2574.
- Kim, J., Fisher III, J.W., Yezzi, A. Çetin, M., Willsky, A.S. (2006). A non parametric statistical method for image segmentation using information theory and curve evolution, *IEEE Trans. Image Processing*, 14(10), 1486-1502.
- Kittler, J., Hatef, M., Duin, R.P.W. (1996). Combining Classifiers. *Proc. 13th. Int. Conf. on Pattern Recognition*, 2, 897-901.
- Kittler, J., Hatef, M., Duin, R.P.W., Matas, J. (1998). On Combining Classifiers, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20 (3), 226-239.
- Kohonen, T. (1989). *Self-Organization and Associative Memory*. Springer-Verlag, Berlin.
- Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE*, 78(9), 1464-1480.
- Kohonen, T. (1997). *Self-Organizing Maps*. Springer, Berlin.
- Koschan, A., Abidi, M. (2008). *Digital Color Image Processing*, Wiley, Chichester.
- Krishnapuram, R., Freg, C.P. (1992). Fitting an unknown number of lines and planes to image data through compatible cluster merging, *Pattern Recognition*, 25, 385-400.
- Kumar, S., Ghosh, J., Crawford, M.M. (2001). Best-bases feature extraction for pairwise classification of hyperspectral data. *IEEE Trans. Geoscience and Remote Sensing*, 39(7), 1368-1379.

- Kumar, A.S., Majumder, K.L. (2001). Information fusion in tree classifiers. *Int. J. Remote Sensing*, 22(5), 861-869.
- Kumar, S., Ghosh, J., Crawford, M.M. (2002). Hierarchical Fusion of Multiple Classifiers for Hyperspectral Data Analysis. *Pattern Analysis and Applications*, 5, 210-220.
- Kuncheva, L.I. (2003). "Fuzzy" vs "non-fuzzy" in combining classifiers designed by boosting. *IEEE Trans. Fuzzy Systems*, 11, 729-741.
- Kuncheva, L.I. (2004). *Combining Pattern Classifiers: Methods and Algorithms*, Wiley.
- Kupinski, M.A., Gigger, M.L. (1998). Automated seeded lesion segmentation on digital mammograms, *IEEE Trans. Medical Imaging*, 17, 510-517.
- LANDSAT Thematic Mapper (2009) [http://edc.usgs.gov/guides/landsat\\_tm.html](http://edc.usgs.gov/guides/landsat_tm.html)
- Lepisto, L, Kunttu, I., Autio, J., Visa, A. (2003). Classification of non-homegenous texture images by combining classifiers. In *Proc. IEEE Int. Conf. Image Processing*, I, 981-984, Barcelona, Spain.
- Lepisto, L, Kunttu, I. Autio, J., Visa, A. (2005). Classification of Natural Images using supervised and Unsupervised Classifier Combinations. F. Roli and S. Vitulano (Eds.). *Lecture Notes in Computer Science*, Springer-Verlag, Berlin, 3617, 770-777.
- Li, H., Liu, G., Zhang, Z. (2006). A new texture generation model based on Pseudo-DCT coefficients, *IEEE Trans. Image Processing*, 15(5), 1300-1312.
- Li, J., Najmi, A., Gray, R.M. (2000). Image classification by a two-dimensional hidden Markov model, *IEEE Trans. Signal Processing*, 48(2), 517-533.
- Liew, A.W.C., Leung, S.H., Lau, W.H. (2003). Segmentation of color images by spatial fuzzy clustering, *IEEE Trans. Fuzzy Systems*, 11, 542-549.
- Lillo, A., Motta, G., Storer, J. A. (2007). Supervised segmentation based on texture signatures extracted in the Frequency Domain, in: J. Martí, J.M. Benedí. A.M. Mendonça, J. Serrat (Eds.), *Pattern Recognition and Image Analysis*, Springer-Verlag, Berlin, LNCS 4477, Part I, 89-96.
- Lim, Y.W., Lee, S.U. (1990). On the color image segmentation algorithm based on the thresholding and the fuzzy c-means technique, *Pattern Recognition*, 23(9), 935-952.

- Linde, Y., Buzo, A., Gray, R.M. (1980). An Algorithm for Vector Quantization Design. IEEE Trans. Communications, 28(1), 84-95.
- Lloyd, S.P. (1957). Least squares quantization in PCM's, Bell Telephone Laboratories Paper, Murray Hill, NJ.
- Lo, Y.S., Pei, S.C. (1999). Color image segmentation using local histogram and self-organizing of Kohonen feature map, in: Proc. Int. Conf. on Image Processing, Kobe, Japan, 232-239.
- Luo, J., Gray, R.T., Lee, H.C. (1998). Incorporation of derivative priors in adaptive Bayesian color image segmentation, in: Proc. Int. Conf. Image Processing, Chicago, 3, 780-784.
- Mandallaz, D., Ye, R. (1999). Forest inventory with optimal two-phase, two-stage sampling schemes based on the anticipated variance, Canadian Journal of Forest Research, 29, 1691-1708.
- Manjunath, B.S., Ma, W.Y. (1996). Texture features for browsing and retrieval of image data, IEEE Trans. Pattern Analysis and Machine Intell., 18(8), 837-842.
- Mas, J.F., Flores, J.J. (2008). The application of artificial neural networks to the analysis of remotely sensed data, Int. J. Remote Sensing, 29(3), 617-663.
- MATLAB (2009). The Matworks <http://www.mathworks.com/>
- Melgani, F., Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines, IEEE Trans. Geosci. Remote Sensing, 42(8), 1178–1790.
- Mohaddamzadeh A., Bourbakis, N. (1997). A fuzzy region growing approach for segmentation of color images, Pattern Recognition, 30(6), 867-881.
- Montes, F., Hernández, M.J., Cañellas, I. (2005). A geostatistical aproach to cork production sampling estimation in Quercus suber L. forests, Canadian Journal of Forest Research, 35, 2787-2796.
- Montiel, E., Aguado, A.S., Nixon, M.S. (2005). Texture classification via conditional histograms, Pattern Recognition Letters, 26, 1740-1751.

- Muñoz-Marí, J., Bruzzone, L., Camps-Valls, G. (2007). A Support Vector Domain Description Approach to Supervised Classification of Remote Sensing Images, *IEEE Trans. Geosci. Remote Sens.*, 45(8), 2683–2692.
- Newsam, S., Wang, L. Bhagavathy, S., Manjunath, B.S. (2004). Using Texture to analyze and Manage Large Collections of Remote Sensed image and Video Data, *Applied Optics*, 43(2), 210-217.
- Özden, M., Polat, E. (2007). A color image segmentation approach for content-based image retrieval, *Pattern Recognition*, 40(4), 1318-132.
- Pajares, G. (2006). A Hopfield Neural Network for Image Change Detection. *IEEE Trans. Neural Networks*, 17(5), 1250-1264.
- Pajares, G., Cruz, J.M. (2007). *Visión por Computador: Imágenes Digitales y Aplicaciones*, RA-MA, Madrid.
- Pajares, G., Cruz, J. M., Moreno, V. (2001). Clasificación de texturas mediante redes neuronales, *Ingeniería Civil. Centro de Estudios y Experimentación de Obras Públicas (CEDEX).*- Ministerio de Fomento, 123, 61-69.
- Pajares, G., Cruz, J.M. (2002). Clasificación de Texturas Naturales mediante K-Means, *Revista Electrónica de Visión por Computador*, (<http://revc.uab.es/revista/06/>, ISSN 1575-5258), no. 6, 1-18, Ed. Centre de Visió per Computador, Universidad Autònoma, Barcelona.
- Pajares, G., Cruz, J. M., Moreno, V. (2002). Clasificación de texturas naturales mediante agrupamiento borroso. *Ingeniería Civil. Centro de Estudios y Experimentación de Obras Públicas (CEDEX).*- Ministerio de Fomento, 127, 83-89.
- Pajares, G., Tellaeche, A., BurgosArtizzu, X., Ribeiro, A. (2007) Design of a computer vision system for a differential spraying operation in precision agriculture through Hebbian learning. *IET Computer Vision*, 1 (3-4), 93-99.
- Palm, C., Lehmann, T.M. (2002). Classification of color textures by Gabor filtering, *Machine Graphics and Vision*, 11(2/3), 195-219.
- Pandya, A.S., Macy, R.B. (1996). *Pattern Recognition with Neural Networks in C++*. CRC Press and IEEE Press, Boca Raton, FL, 410pp.

- Paola, J.D., Schowengerdt, R.A. (1995). A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery, *Int. J. Remote Sensing*, 16(16), 3033-3058.
- Pappas, T.N. (1992). An adaptive clustering algorithm for image segmentation, *IEEE Trans. Signal Processing*, 40(4), 901-914.
- Parzen, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.*, 33, 1065-1076.
- PCI Geomatics (2009). <http://www.pci.on.ca/>
- Permuter, H., Francos, J., Jermyn, I. (2006). A study of Gaussian mixture models of color and texture features for image classification and segmentation, *Pattern Recognition*, 39, 695-706.
- Petrakos, M., Benediktsson, J.A., Kanellopoulos, I. (2001). The effect of classifier agreement on the accuracy of the combined classifier in decision level fusion. *IEEE Trans. Geosci. Remote Sens.*, 39(11), 2539–2546.
- Pi, M.H., Tong, C.S., Choy, S.K., Zhang, H. (2006). A fast and effective model for wavelet subband histograms and its applications in texture image retrieval, *IEEE Trans. Image Processing*, 15(10), 3078-3088.
- Porat, M., Zeevi, Y.Y. (1989). Localized texture processing in vision: analysis and synthesis in the Gaborian space, *IEEE Transactions on Biomedical Engineering*, 36(1), 115-129.
- Precioso, F., Barlaud, M., Blu, T., Unser, M. (2005). Robust real time segmentation of images and videos using a smooth-spline snake-based algorithm, *IEEE Trans. Image Processing*, 14(7), 910-924.
- Proespacio (2009). Agrupación de empresas. <http://www.proespacio.org/>
- Puig, D., García, M.A. (2006). Automatic texture feature selection for image pixel classification, *Pattern Recognition*, 39(11), 1996-2009.
- Pyun, K., Won, C.S., Lim, J., Gray, R.M. (2002). Texture classification based on multiple Gauss mixture vector quantizer, in: *Proc. of ICME*, 501-504.
- Randen, T., Husøy, J.H. (1999). Filtering for Texture Classification: A Comparative Study, *IEEE Trans. Pattern Analysis Machine Intell.*, 21(4), 291-310.

- Ríos-Insúa, S., Bielza, C., Mateos, A. (2002) Fundamentos de los Sistemas de Ayuda a la Decisión. RA-MA.
- Robinson, G.S. (1977). Color edge detection, *Optical Engineering*, 16(5), 479-484.
- Rud, R., Shoshany, M., Alchanatis, V., Cohen, Y. (2006). Application of spectral features' ratios for improving classification in partially calibrated hyperspectral imagery: a case study of separating Mediterranean vegetation species, *Journal Real-Time Image Processing*, 1, 143-152.
- SAR (2009). Synthetic Aperture Radar, Jet Propulsion Laboratory. <http://southport.jpl.nasa.gov/>
- Seong, J.C., Ustry, E.L. (2001). Fuzzy image classification for continental scale multitemporal NDVI series images using invariant pixels and an image stratification method, *Photogrammetric Engineering and Remote Sensing*, 67(3), 287-294.
- SIC, Satellite Image Corporation (2009). <http://www.satimagingcorp.com/>.
- Singh, S., Haddon, J., Markou, M. (2001). Nearest-neighbour classifiers in natural scene analysis, *Pattern Recognition*, 34, 1601-1612.
- Smith, G., Burns, I. (1997). Measuring texture classification algorithms, *Pattern Recognition Letters*, 18, 1495-1501.
- Tellaeche, A., BurgosArtizzu, X.P., Pajares, G., Ribeiro, A. (2008). A vision-based method for weeds identification through the Bayesian decision theory. *Pattern Recognition*, 41(2), 521-530.
- Tkalcic, M., Tasic, J.F. (2003). Colour spaces-perceptual, historical and applicational background, in: *EUROCON 2003, Computer as a Tool, The IEEE Region 8*, vol. 1, pp. 304-308.
- Tottrup, C. (2007). Forest and Land Cover Mapping in a Tropical Highland Region, *Photogrammetric Engineering and Remote Sensing*, 73(9), 1057-1066.
- Trauwaert, E., Kaufman, L., Rousseeuw, P. (1991). Fuzzy Clustering Algorithms Based on the Maximum Likelihood Principle, *Fuzzy Sets and Systems*, 42, 213-227.



- Trias-Sanz, R., Stamon, G., Louchet, J. (2008). Using colour, texture, and hierarchical segmentation for high-resolution remote sensing, *ISPRS Journal of Photogrammetry & Remote Sensing*, 63, 156-168.
- Tso, B.C.K., Mather, P.M. (1999). Classification of multisource remote sensing imagery using a genetic algorithm and Markov random fields, *IEEE Trans. Remote Sensing*, 37(3), 1255-1260.
- Unser, M. (1995). Texture classification and segmentation using wavelet frames, *IEEE Trans. Image Processing*, 4(11), 1549-1560.
- Valdovinos, R.M., Sánchez, J.S., Barandela, R. (2005). Dynamic and Static weighting in classifier fusion, in: J.S. Marques, N. Pérez de la Blanca and P. Pina (Eds.), *Pattern Recognition and Image Analysis*, Springer-Verlag, Berlin, LNCS 3523, pp. 59-66.
- Valdovinos, R.M., Sánchez, J.S. (2007). Performance Analysis of Classifier Ensembles: Neural Networks Versus Nearest Neighbor Rule. (J. Martí, J.M. Benedí. A.M. Mendonça and J. Serrat (Eds.)). *Pattern Recognition and Image Analysis. Lecture Notes in Computer Science*, Springer-Verlag, Berlin, 4477, Part I, 105-112.
- Vazquez, E., Baldrich, R. Vazquez, J., Vanrell, M. (2007). Topological Histogram Reduction towards colour segmentation, in: J. Martí, J.M. Benedí. A.M. Mendonça and J. Serrat (Eds.), *Pattern Recognition and Image Analysis*, Springer-Verlag, Berlin, LNCS 4477, Part I, 55-62.
- Volkovich, Z., Barzily, Z., Morozensky, L. (2008). A statistical model of cluster stability. *Pattern Recognition*, 41(7), 2174-2188.
- Wagner, T. (1999). Texture Analysis, in: B. Jahne, H. Haussecker, P. Geisser (Eds.), *Handbook of Computer Vision and Applications*, Academic Press, San Diego, cap. 2 (Signal Processing and Pattern Recognition).
- Wan, T., Canagarajah, N., Achim, A. (2007). Multiscale Color-Texture image segmentation with adaptive Region Merging, In *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP08)*, 1, 1-1213-1216.

- Wang, Z., Boesch, R. (2007). Color- and Texture-Based Image Segmentation for Improved Forest Delineation, *IEEE Trans. Geoscience and Remote Sensing*, 45(10), 3055–3062.
- Wang, W., Fenton, N. (2006). Risk and confidence analysis for fuzzy multicriteria decision making. *Knowledge Based Systems*, 19, 430-437.
- Wangenheim, A., Bertoldi, R.F., Abdala, D.D., Richter, M.M. (2007). Color image segmentation guided by a color gradient network, *Pattern Recognition Letters*, 28, 1795-1803.
- Wilkinson, G.G. (2005). Results and implications of a Study of Fiteen Years of Satellite Image Classification Experiments, *IEEE Trans. Geoscience Remote Sensing*, 45(3), 433-440.
- Wilkinson, G.G., Fierens, F., Kanellopoulos, I. (1995). Integration of neural and statistical approaches in spatial data classification. *Geograph. Syst.*, 32, 1–20.
- Wozniak, M., Jackowski, K. (2008). Fusers Based on Classifier Response and discriminant Function-Comparative Study. (E. Corchado, A. Abraham and W. Pedrycz (Eds.). *Hybrid Artificial Intelligent Systems. Lecture Notes in Artificial Intelligent*, Springer-Verlag, Berlin, 5271, 361-368.
- Xia, Y., Feng, D., Wang, T., Zhao, R., Zhang, Y. (2007). Image segmentation by clustering of spatial patterns, *Pattern Recognition Letters*, 28, 1548-1555.
- Xie, J., Tsui, H.T. (2004). Image Segmentation based on maximum-likelihood estimation and optimum entropy-distribution (MLE-OED), *Pattern Recognition Letters*, 25(10), 1133-1141.
- Yang, A.Y., Wright, J., Ma, Y., Sastry, S.S. (2008). Unsupervised segmentation of natural images via lossy data compression, *Computer Vision and Image Understanding*, 110, 212-225.
- Yu, H., Li, M., Zhang, H.J., Feng, J. (2002). Color texture moments for content-based image retrieval, in: *Proc. International Conference on Image Processing*, 3, 24-28.
- Zadeh L.A. (1975). The concept of a linguistic variable and its application to approximate reasoning. *Inform. Sci.* 8 (1975) 199-249(I), 301-357(II).

- Zhang, J., Foody, G.M. (1998). A fuzzy classification of sub-urban land cover from remotely sensed imagery, *Int. J. Remote Sensing*, 19(14), 2721-2738.
- Zhang, Y., Wang, R. (2004). A combined method for texture analysis and its application, in: F. Bubak, M. (Eds.), *Lecture Notes in Computer Science*, 3036, 413-416.
- Zimmermann, H.J. (1991). *Fuzzy set theory and its applications*, Kluwer Academic Publishers, Norwell, MA.